ANIMAL HEALTH INSTITUTE

Representing manufacturers of animal health products

Richard A. Carnevale, VMD
Vice President, Regulatory, Scientific and International Affairs

April 19, 2004

Division of Dockets Management
Food and Drug Administration
5630 Fishers Lane, Room 1061 (HFA-305)
Rockville, Maryland 20852

Re: Comments to FDA Docket No. 98D-1146, Draft Guidance for Industry #152 "Evaluating the Safety of Antimicrobial New Animal Drugs with Regard to Their Microbiological Effects on Bacteria of Human Health Concern"

On July 25, 2003, the Animal Health Institute submitted for consideration an alternative approach to the CVM guidance document #152 to assess the microbiological safety of antimicrobial agents used in food producing animals.

Further refinements have been added to the document and we are please to submit for consideration, "Human Health Impacts of Animal Antibiotics: A Guide to Health Risk and Benefit Analysis using the Rapid Risk Rating Technique."

We believe this model provides a more thorough approach to evaluating the food safety of antimicrobial products from the potential selection of resistant organisms. Our members might prefer to utilize this model when developing the human safety package for an NADA submission since it provides a more quantitative approach to assessing risk while maintaining the concepts outlined in Guidance document #152.

We sincerely hope that FDA will give serious consideration to this submission and advise AHI accordingly.

Sincerely,

Richard A. Carnevale, VMD

Enclosure

98D-1146

C 106

1325 G Street, NW ■ Suite 700 ■ Washington, D.C. 20005-3104
Telephone (202) 637-2440 ■ Fax (202) 393-1667
www.ahi.org

Human Health Impacts of Animal Antibiotics

A Guide to Health Risk and Benefit Analysis Using the Rapid Risk Rating Technique

Revised
April 15, 2004

Prepared by:

The Animal Health Institute
1325 G Street
Washington, DC

and

Cox Associates
503 Franklin Street
Denver, Colorado

Outline of Contents

EXECUTIVE SUMMARY

At the beginning of 2004, human health risk management of animal antibiotic use has reached a crucial juncture for public health officials and risk management policy-makers worldwide. There is an opportunity to use Risk Analysis in regulatory and public health areas as the overarching process that will bring divergent viewpoints and complex real-world interactions into focus, to allow for appropriate risk management decisions to be made and communicated.

This report reviews the logic, process, and methods of quantitative risk analysis for food borne bacterial and antimicrobial resistance hazards, as well as several proposed qualitative scoring approaches. It shows that a quantitative Rapid Risk Rating Technique (RRRT) based on top-down estimation and multiplication of Exposure, Dose-Response, and Consequence factors, as suggested by WHO (2003), is at least as simple as qualitative rating approaches and gives more useful, data-driven and meaningful results. Several examples illustrate the approach. Guidance is offered for how best to carry out the steps in a risk assessment and risk analysis process for animal antibiotics to identify risk management actions that are most likely to protect human health.

BACKGROUND

Risk assessment has long been endorsed by national governments and international organizations as a means for authorities to use science to prioritize the use of limited resources to concentrate efforts against those hazards of most concern to public health. In the United States, the National Academy of Sciences developed criteria in a 1983 study that was partially sponsored by the Food and Drug Administration. The 1983 NAS report set forth the essential elements of risk assessment which are in common use today. These serve as the basis for the subsequent work of Codex, WHO and FAO in reviewing chemical and food residue risks and, more recently, in assessing microbial food-borne hazards.

In 1997, the Codex Alimentarius Commission (CAC) adopted the following "Statements of Principle Relating to the Role of Food Safety And Risk Assessment (ALINORM 97/33)":

1. Health and safety aspects of Codex decisions and recommendations should be based on a risk assessment, as appropriate to the circumstances.

2. Food safety risk assessment should be soundly based on science, should incorporate the four steps of the risk assessment process [i.e., hazard identification, exposure assessment, exposure-response modeling, and risk characterization], and should be documented in a transparent manner.

3. There should be a functional separation of risk assessment and risk management, while recognizing that some interactions are essential for a pragmatic approach.

4. Risk assessments should use available quantitative information to the greatest extent possible and risk characterizations should be presented in a readily understandable and useful form.

In March 1999, the WHO/FAO held a consultation in Geneva, Switzerland on the application of risk assessment to microbiological hazards in foods. This was followed by several other consultations to develop risk assessment documents on specific food borne pathogens such as Listeria, Campylobacter, and Vibrio spp. The consultation served to advise member countries and the CAC on the use of this methodology in setting food standards important to public health and international food trade.

In June 1999, the CAC adopted "Principles and Guidelines for the Conduct of Microbiological Risk Assessment" (CAC/GL-30 1999), which set forth the scope, definitions and guidelines for the conduct of quantitative or qualitative risk assessment to microbial hazards. The definitions and guidelines follow very closely the principles first elaborated by the National Academy of Sciences 20 years ago.

The Office of International Epizootics Ad Hoc Group on Antimicrobial Resistance elaborated principles for risk analysis tailored to address antimicrobial resistance in animals. The OIE principles are similar in many respects to those elaborated by the Codex and NAS

for microbial risk assessment but may not be as applicable to food safety since the OIE is the organization primarily involved with setting standards for the control and prevention of global animal diseases. Evaluation of food safety risks involves a farm to table evaluation with consideration of factors well beyond the farm. Key recommendations from the OIE to effectively manage antimicrobial resistance risk issues include the following (http://www.oie.int/eng/publicat/rt/2003/VOSE.PDF):

1. "Risk analysis should be conducted in an objective and defensible manner
2. The risk analysis process should be transparent and consistent - risk analysis should be conducted as an iterative and continuous process
3. Risk management and risk assessment functions should be kept separate to ensure the independence of decision-making and evaluation of the risk
4. Risk management should be conducted in reference to a policy framework setting out the domain of the regulator and the range of risk reduction actions that may be considered
5. The risk assessment should be based on sound science and conducted according to a strategy established by the risk managers in co-operation with the risk assessors
6. Risk assessment requires a multidisciplinary team and should be conducted in broad consultation with available scientific expertise
7. Qualitative risk assessment should always be undertaken, and provides information on whether progression to full quantitative risk assessment is feasible and/or necessary
8. Risk assessment of antimicrobial resistance issues requires very specific, technical skills that may not be available to developing countries. The OIE and its Member Countries should work towards helping these countries to develop or access these skills, to ensure that risk assessment itself does not become a barrier to trade
9. Communication between managers, assessors and stakeholders is essential. Effort should be made to establish such communication early in the process, to allow opportunity for responses, and should be continued throughout the risk analysis process."

In 1994, the WTO Agreement on the Application of Sanitary and Phyto-Sanitary Measures (SPS Agreement) established science as the key basis for measures by WTO Members aimed at protecting human health and animal or plant life or health. The WTO SPS Agreement recognizes the role of science, harmonization, and international standards-setting bodies in formulating public health and food safety measures. Article 2.2 of the SPS Agreement states that such measures should be "based on scientific principles" and on "sufficient scientific evidence." In addition, Article 5.1 provides that WTO Members "shall ensure that their sanitary and phyto-sanitary measures are based on an assessment, as appropriate to the circumstances, of the risks to human, animal or plant life or health, taking into account risk assessment techniques developed by the relevant international organizations."

This document presents principles of risk analysis and risk assessment for food-borne pathogens, including both susceptible and resistant strains. It builds on the earlier efforts just

described, and seeks to distill the most sound and useful principles for supporting improved risk management decision-making with data-based, scientifically valid analysis, as required by the WTO.

Objectives

This guideline was developed to support risk analysis for the use of antimicrobial agents in food animals and their potential impact to the public health. It builds on previous Codex guidelines for microbiologic risk assessment and specifically to guidelines relating to antimicrobial resistance published by the OIE, US FDA-CVM, and Australia.

This risk evaluation guideline is based on WHO/FAO guidelines, but also considers the best practices among individual country guidelines and approaches to risk assessment. It emphasizes the transparency, objectivity, and logic of the best risk analysis guidelines, as well as the practicality of using them with existing data. It seeks to provide principles useful for qualitative, semi-quantitative and fully quantitative risk analyses.

Positioning of this guideline in the risk analysis process

Risk analysis has been defined in Codex as "The process composed of risk assessment, risk management and risk communication". It describes a complete process for addressing a risk issue. It encompasses assessing and managing the risk together with all the appropriate communication between risk assessors, stakeholders and risk managers.

This guideline defines procedures for preparing risk assessment reports that would be consistent with internationally-accepted practice for risk assessment and thus to support a national regulatory authority in assessing and managing risks. The risks considered include probability of the loss of benefit of antimicrobial therapy in humans due to acquired resistance (resistant bacterium or resistant determinants) through the use of a specific antimicrobial agents in animals. This guideline also incorporates potential human health benefits of antimicrobial feed additives, such as, improved animal health, food production, food security, and improved food safety.

OIE states that "qualitative risk assessment should always be undertaken, and provides information on whether progression to full quantitative risk assessment is feasible and/or necessary". Thus, qualitative risk assessment can play a valuable screening and prioritization role. This document therefore discusses principles and procedures for qualitative risk assessment as well as for quantitative risk assessment.

Methods for conducting quantitative risk assessments have been extensively developed, and many detailed technical methods and principles of study design and data analysis are now available to support the successful execution of microbial risk assessments and antimicrobial risk assessments in food safety. Appendix A outlines relevant technical approaches and methods that can be useful in achieving the goals of risk analysis.

RISK ANALYSIS

DEFINITION OF RISK ANALYSIS

Risk analysis provides methods, principles, and high-level procedures for using scientific data to assess and compare the probable consequences of exposures to different hazards (i.e., sources of risk) and to rationally evaluate and choose among alternative risk management decision options. It is often divided into stages of *risk assessment, risk management,* and *risk communication,* organized as an iterative process. Table 1 summarizes several traditionally defined steps in this process. Risk analysis quantifies the probable human health consequences, both positive and negative, and other (e.g., animal health, environmental) consequences of alternative risk management decisions.

Health risk *assessment* estimates the health risks to individuals, groups (e.g., the old, the young, or the immune-compromised), and entire populations from hazardous exposures and from the decisions or activities that create them. Health risks are measured by the changes in probabilities and magnitudes (or in frequencies and severities) of adverse health effects caused by exposures. Individual risks of sporadic illnesses may be expressed as *expected numbers of additional adverse health effects per capita-year*, by severity category (e.g., mild, moderate, severe, or fatal; see e.g., Buzby, et al., 1996). Population risks sum individual risks over all person-years in the population. They may be expressed as *expected numbers of additional adverse health effects per year* of each type or clinical severity category occurring in the population. Population risks may also be further described by identifying sub-populations with especially high individual risks from exposure.

Technical Note: Poisson Approximation. Use of expected number of events per year to quantify risk is justified for sporadic illnesses that occur independently or with only weak statistical dependence in large populations under the conditions of the Poisson or compound Poisson approximations. The expected number of cases per year then determines the full probability distribution of the total number of illnesses per year. For exact mathematical results, see http://citeseer.nj.nec.com/372423.html; Barbour, 2000.

Following the National Academy of Sciences framework for risk analysis, the US FDA, CDC and USDA defined *risk assessment* as a process that "consists of the following steps: hazard identification, exposure assessment, hazard characterization (dose-response), and risk characterization" (http://www.foodsafety.gov/~dms/lmriskgl.html). Dose-response assessment is defined as "The determination of the relationship between the magnitude of exposure and the magnitude and/or frequency of adverse effects." Similar concepts have been adopted internationally in WHO/FAO guidelines and OIE guidelines (http://www.oie.int/eng/publicat/rt/2003/VOSE.PDF).

The main goal of risk assessment is to produce information to improve risk management decisions. It does this by *identifying causal relations between alternative risk management decisions and their probable total human health consequences* (including health benefits as well as health risks) and by identifying those decisions that make preferred outcomes more likely. Unlike informal expert decision analysis or judgment-based methods, health risk assessment uses explicit analytic (e.g., biomathematical, statistical, or simulation) models of the causal relations between actions and their probable health effects. In general, quantitative risk assessment applies specialized models and methods to quantify likely exposures and the frequencies and severities of their resulting health consequences.

TABLE 1: Steps in Traditional Risk Assessment Framework

Step	Purpose and Description	Relevant information and techniques	Types of data that could be included for feed additive risk assessment
Hazard identification	Identify potential sources of harm or loss. These sources are called *hazards*. Hazard identification identifies possible adverse health effects of activities or exposures and possible causes of observed adverse effects.	• Human data: Epidemiology, clinical and public health statistics. • Animal tests and bioassays • *In vitro* tests • Structure-activity patterns, molecular modeling, pattern recognition and statistical classification techniques	• Bacteria that have acquired resistance or resistance determinants due to the particular use of an antimicrobial in animals • Activity of drug, spectrum, bacteria of concern, importance of drug to human medicine
Exposure assessment	Quantify the number of people receiving various levels or intensities of exposure to a hazard over time. Relevant exposure metrics may depend on dose-response relations.	Environmental fate and transport models, possibly summed over multiple media (paths) and sources Studies of human activity patterns Biological monitoring of exposed individuals and receptors	• Animal use patterns, resistance mechanisms, genetic transfer, pharmaocokinetic data for gut activity • Microbial contamination during processing, cooking of food; consumption patterns and demographics of human populations, etc. • Dose of bacteria causing disease; human transmission • Resistance monitoring results
Quantitative exposure-response and dose-response modeling	Quantify the magnitude of risk created by exposure of a target to a hazard. Characterize the probable frequency and severity of adverse health outcomes or losses caused by exposure to the hazard.	A quantitative risk assessment (QRA) runs multiple exposure scenarios through a *dose-response model* to predict likely health impacts. Statistical, simulation, or biomathematical models of biological processes are used to quantify dose-response relations.	• Foodborne disease antibiotic treatment options and outcomes; importance to human medicine • Resistance monitoring results • Perceived future of the drug
Risk characterization	Combine estimated probabilities and severities of health harm (adverse consequences), together with indications of uncertainty or confidence, to create an overall summary and presentation of risk.	Monte Carlo simulation calculates risks by sampling multiple scenarios. Risk profiles, probability distributions, and trade-off and sensitivity analyses display risk, uncertainty, and variability.	No example provided
Risk communication	Deals with how to present risk information to stakeholders. Considers how different types of recipients perceive risks and internalize/act on messages about them, in deciding what messages to send via what media.	Psychological theories and models and behavioral/experimental findings on risk perception and effective risk communication.	No example provided
Risk management decision-making	Decide what actions to take to control risks and hazards – i.e., accept, ban, abate, monitor, further research, reduce, transfer, share, mitigate, or compensate.	Risk-cost-benefit analysis, formal decision analysis for groups and individuals, risk quantification and comparison	• Summary of benefits to humans and animals from use of the product • Justification of risk management actions for effectiveness, cost, value, etc.

Health risk *management* applies principles for choosing among alternative decisions, policies or actions that affect exposures, health risks, or their consequences. Risk management is often viewed as a process that takes scientific information obtained from risk assessment as input and that recommends choices of risk management actions as output. Alternative risk management approaches may include risk acceptance, prevention or avoidance (e.g., by reduction of microbial loads during processing or food preparation), mitigation of consequences (e.g., by appropriate clinical screening, diagnosis, and prescription procedures), transfer (e.g., health insurance) or compensation.

Health risk *communication* characterizes and presents information about health risks and uncertainties to decision-makers and stakeholders. Risk assessment and risk communication should support effective risk management decision-making by providing the scientific information needed to compare alternative risk management interventions in terms of their probable impacts on exposures and the resulting frequency and severity of adverse health effects. If animal antibiotics reduce the frequency and severity of some adverse human health effects, then these impacts should be included as part of the complete risk assessment and communication package and should be taken into account in risk management decision-making.

PURPOSES OF RISK ANALYSIS

The primary purpose of health risk analysis is to support improved risk management decision-making. By definition, "better" risk management decisions are more likely to produce preferred consequences, i.e., fewer illnesses, mortalities, illness-days, and treatment failures per person-year. Health risk analysis also provides a framework for rational deliberation, conflict resolution, policy-making, and international harmonization about human health risks of commercial activities. It allows better-informed and more effective regulation of the production, distribution, preparation, and use of antimicrobials in food animals than approaches that are not driven by analysis of probable consequences of alternative decisions. Risk analysis also provides a framework for predicting how such activities interact with human behaviors – e.g., consumer or food worker behaviors in food handling and kitchen hygiene; physician decisions about what tests and treatments to prescribe when to which patients; and patient decisions about seeking and complying with physician instructions on antibiotic use – in determining the frequencies and magnitudes of adverse health outcomes.

The risk management decision alternatives to be evaluated by risk assessment are often of the following types:

- *Status quo option:* Do not take actions that will change current exposure patterns
- *Restriction or ban:* Take action to reduce current exposures to hazards. Examples include training and education programs, monitoring and enforcement activities, and HACCP programs.
- *Approval of new product or process:* Take action (e.g., approve a new product, use, or product line extension) that may increase or modify current exposure patterns

Different data are typically available for evaluating these three types of options, with changes in uses of products that have already been used for many years often having the most data, while approvals of new products must rely more on plausible worst-case assumptions, models and/or analogies to existing products. However, the same general logical assessment process applies to all types of risk management options.

The main value and purpose of risk assessment in such cases is usually to quantify and compare the probable human health risks (i.e., expected change in the expected absolute or relative number and/or severity of foodborne illness cases per year in exposed populations) for *each* risk management decision option considered, *conditioned* on whatever information is available about it. Computational-statistical, mathematical and probability modeling, and computer simulation methods enable risk assessors to constrain and estimate human health risks and uncertainties quantitatively from realistic (incomplete, imprecise, inaccurate and perhaps inconsistent and incorrect) measurements and data. A pragmatic risk analysis perspective is that decisions can often be informed and improved even by imperfect measurements and incomplete facts, knowledge and data, if they provide some statistical information about probable human health consequences of risk management alternatives. This perspective is sometimes formalized in "value-of-information" (VoI) and sensitivity analysis calculations. It allows risk assessment to deliver useful results about probable risks and remaining uncertainties based on currently available empirical information, while also showing the potential for these results to change as further information is collected and indicating which information is likely to lead to the greatest changes. This information, in turn, is what risk management decision-makers need to make rational interim decisions and to identify what new empirical information would be required to justify future changes.

DESIRED OUTPUTS OF RISK ANALYSIS

A successful risk analysis shows the estimated changes in frequencies and magnitudes of human heath consequences caused by different risk management decision options. It also uses confidence intervals and other qualitative and/or quantitative displays to show uncertainties about the human health consequences of different decisions. It identifies a subset of one or more decision options leading to preferred probability distributions of health risks. Thus, a successfully completed risk analysis should allow a decision-maker to answer the following questions for each risk management decision alternative being evaluated or compared:

- *What probable change in human health risk would be caused by each risk management intervention?* If the risk management decision option or action being assessed is implemented, how will the probable adverse human health effects (e.g., expected numbers of mild, moderate, severe, and fatal illnesses per year; expected numbers of illness-days and, if desired, quality-adjusted life-years (QALYs) lost per year) change in the whole population and in sub-populations with distinct risks?

- *How certain is the change in human health risk that would be caused by each risk management action?* Instead of a single value, i.e., a "point estimate" of risk, uncertain risks are characterized by intervals or probability distributions indicating how closely the change in human health risk caused by a proposed risk management intervention can be predicted. There are several technical options for expressing uncertainty around point estimates (e.g., plausible upper and lower bounds or confidence limits, coefficients of variation, confidence intervals, tolerance intervals, prediction intervals, Bayesian probability intervals, Bayesian posterior distributions, etc.) More elaborate uncertainty displays (e.g., confidence contours for the joint distribution of frequency and severity components of risk) are available for specialists. The essential information to provide about uncertainty in any risk assessment is how large or how small the true risks might be, consistent with the data and with the specified assumptions of the risk assessment. Point estimates that are "best" with respect to various technical statistical criteria will typically fall between these extremes.

  Technical note: Statistical point estimates and interval estimates. Many criteria have been used to define and identify "best" point estimates in risk models, e.g., maximum likelihood estimates (MLE), maximum a posteriori (MAP) Bayesian estimates, maximum entropy, minimum description length, least squares, minimum absolute deviation, minimum expected loss (for various loss functions). While these criteria have led to useful theory and algorithms for estimating the parameters of risk models, *none* of them is satisfactory as the sole output from a risk assessment. *It is essential to provide intervals or probability distributions around any point estimate of risk* to inform the users of a risk assessment about the full range of risks that might be caused by a risk management intervention. This principle applies to qualitative and fuzzy risk ratings as well. If a point estimate of a risk is "High", then some indication must be given of how certain this value is and of how compatible the frequency and severity components of the risk are with other qualitative labels, such as "Low". A risk assessment that produces a single overall value for risk with no indication of uncertainty should be avoided.

- *What are the key drivers of risks and uncertainties for each option?* The analysis should make clear to the user the main reasons *why* the estimated risk from each decision option is as high or low as it is. Are the results driven mainly by predicted exposure levels, by the responses of sensitive sub-populations, by genetic or epidemiological data that establishes tight constraints on the plausible values, or by other factors? Sensitivity analyses that plot how estimated risks would change as input assumptions and estimates vary within plausible ranges (e.g., within a few standard deviations of their median values) can help to identify visually the combinations of input values that drive the main conclusions and the extent to which these could be changed without changing the comparison of different risk management interventions.

- *Which risk management interventions are undominated?* One risk management intervention *dominates* another if it produces smaller probabilities of exceeding any specified level of adverse consequences per year. For example, if two different interventions lead to different expected numbers of sporadic salmonellosis cases per year (with the actual number being a Poisson random variable), and if the probable health consequences per case (e.g., the number of days of illness of a given severity) is the same for each intervention, then the one giving the smaller expected number of illnesses per year dominates the other. If the expected number of cases per year for each

intervention is uncertain, and if the probability that it exceeds any specified value is smaller for intervention A than for intervention B (for all possible specified values), then A dominates B. Scientific risk assessment can, at most, identify undominated risk management alternatives for risk managers to further assess and choose among, but stops short of being able to recommend an objectively "best" choice among multiple undominated interventions.

EXAMPLE OF RISK ANALYSIS LOGIC AND OUTPUT: THE RRRT FRAMEWORK

Although risk assessment models can be technically sophisticated and detailed, the main logic of health risk assessment is often straightforward and relatively easy to validate, to a useful degree, with empirical data. For example, the population risk of a foodborne illness can often be modeled as a product of factors, as follows:

> Population Risk = (expected number of contaminated servings ingested per year) * (expected illnesses caused per contaminated serving ingested) * (expected illness-days, mortalities, or other adverse consequences per illness) = Expected number of adverse human health consequences per year in the population.

This product model can be abbreviated as:

Risk = (exposure factor)*(dose-response factor)*(consequence factor).

Such a multiplicative, top-down approach has been recommended on methodological grounds (Bailar and Travers 2002; FSRC, 2003) and has long been used by practitioners. Review of past studies and data may be required to estimate and document the values and uncertainty intervals (or approximate probability distributions) for these factors. While doing so may involve detailed data collection, calculations, and modeling, the overall logic is relatively simple and transparent. It can be applied to many direct risks from foodborne pathogens. An intervention that changes one or more of these factors will change the predicted population risk correspondingly. If a risk management intervention simultaneously affects multiple contaminants (e.g., multiple pathogens, or both susceptible and resistant strains of a pathogen), multiple food commodities, and/or subpopulations having distinct exposure-response relations, then summing the above product over all combinations of these multiple components gives the total impact on population risk. This is the basis of the Rapid Risk Rating Technique (RRRT), discussed later.

Table 2 shows an example of the beginning of a risk assessment calculation made in this RRRT framework. Appendix B explains the details (including the symbols used in the first column). For purposes of illustrating a portion of a risk analysis calculations and output, the most important points about the RRRT framework are as follows:

- *Scope is matched to decision option being evaluated:* The *scope* of the risk assessment calculation in Table 2 is to estimate the number of macrolide-resistant *C. jejuni* cases per year that: (a) Might be caused by use of macrolide products (e.g., Tylosin Premix

and Tylosin Soluble) in chickens; and (b) Are severe enough to warrant treatment with erythromycin or another macrolide in current clinical practice, i.e., some potential clinical benefit might be achieved if the treatment is effective. Such a calculation would presumably be relevant for bounding the potential human health benefits (or, equivalently, human health risk reductions) from a risk management intervention that specifically restricts or eliminates the use of macrolides in chickens. It would not be appropriate for evaluating a ban on all animal growth promoters in a certain class (as in Europe) or evaluating introduction of a new product (e.g., a macrolide product line extension). However, it is appropriate for evaluating a risk management intervention that specifically affects macrolide use in chickens. To complete the risk assessment, it would be necessary to carry out similar calculations for other pathogens (e.g., *C. coli*) affected by the risk management intervention being assessed. Rather than pursuing this in detail, an initial rapid screening assessment might simply document as an assumption that the risks from chicken-borne *C. coli* are not greater than those from chicken-borne *C. jejuni*, and use this assumption to bound the additional contribution from *C. coli*.

TABLE 2: Example of a Top-Down Risk Assessment for Macrolide Use in Chickens

Variable	Values, Uncertainty Factors (UF)	Data Sources
EXPOSURE: CURRENT RESISTANT CASES CAUSED BY ANIMAL ANTIBIOTIC USE		
Total current campylobacteriosis cases reported per 100,000 people per year	13.37 cases/100,000 in 2002 for FoodNet surveillance sample, UF ≈ 1	CDC, 2003
Fraction of *C. jejuni* cases that are severe i.e., treatment with antibiotic is indicated	0.00595. (Uncertainty analyzed via sensitivity analysis.)	Buzby, et al., 1996
Average total severe cases per reported severe case	8 (Ranges from 2 for severe cases to 38 for mild cases; UF = **5**)	Mead et al. (1999)
US population, N	292E6 = 2,920 x 100,000 people in US, UF ≈ 1	US Census Bureau
Fraction of severe cases that are *C. jejuni*	0.99 (May be as low as 0.95), UF ≈ 1	CDC DBMD
Fraction of severe *C. jejuni* cases that are caused by chicken products (including cross-contamination of other foods)	0.10, uncertainty factor = 3-**10**, estimate based on competing risk, genetic, epidemiological, and historical data	Appendix B; Stern and Robach, 2003
Fraction of chicken-caused severe cases that are antibiotic-resistant, (1 – s)	0.01 for erythromycin resistance, UF = **2**	CDC, 2000
Resistant severe *C. jejuni* cases per year caused by chicken products = ($\bar{P}$)*(MN)*(1- s) = product of above	**1.84** cases/yr. for macrolides = (13.37E-5) * 0.00595*8*292E6*0.99* 0.10*0.01; UF = 18 (from component UFs of 5, 10, 2)	Product of above.

- *Simple, transparent calculation logic* (Bailar and Travers, 2002). The calculations are based on multiplying a sequence of factors that have been estimated from documented data sources. Thus, if any of the cited values is thought to be inappropriate, or if more recent data become available, the specified values of the factors can be easily updated and the results recalculated.
- *Clearly interpretable output.* The main result in Table 2 is that the number of "Resistant severe *C. jejuni* cases per year caused by chicken products" is estimated to be 1.84 cases per year. This point estimate is accompanied by an uncertainty factor (explained below) of about 18, corresponding to a subjective Bayesian 95% probability

interval of [1.84/18, 1.84*18] = [0.1, 33] cases per year. This intermediate result does not yet consider the *consequence* component of risk, i.e., what fraction of these cases will seek medical care, be prescribed a macrolide antibiotic, experience treatment failure due to resistance, and suffer excess days of illness. Nor does it consider the *preventable fraction* of the exposures, i.e., the fraction of macrolide-resistant *C. jejuni*-contaminated chicken servings that would be removed (and presumably replaced by macrolide-susceptible *C. jejuni*-contaminated servings) in the event of a risk management intervention. By organizing the calculations as a multiplicative sequence, however, it becomes possible to stop the calculation part way through, yielding an upper-bound estimate on the final result of the complete exposure, illness, and consequence product calculation (since multiplication by additional fractions can only reduce the current result.) Thus, 1.84 is an upper bound for the point estimate of the preventable number of cases per year that may experience a loss of clinical benefits due to macrolide-resistant *C. jejuni* from macrolide-exposed chickens.

- *Key drivers and sensitivity analyses.* Inspection of the numerical values of the factors in the product calculation shows which ones have the greatest impact on the final results. In conjunction with uncertainty factors indicating how many times too high or too low the estimated values might plausibly be compared to the true values, the estimated point values show what changes in factors might occur as additional information is collected and by how much such changes could increase or decrease the current point estimate of risk. In Table 2, the estimated fraction of cases that are severe enough to potentially benefit from antibiotic therapy is obviously a crucial parameter, as it reduces the overall product by a factor of 0.00595. This value is obtained from Buzby et al., 1996. Increasing or decreasing it to reflect more recent data, when they become available, will increase or decrease estimated risks proportionally.
- *Uncertainty analysis.* Uncertainty factors of about 1 (UF $\approx$ 1) in Table 2 indicate quantities that are known with sufficient precision so that better information about them is not expected to make a significant change in the results. Uncertainty factors greater than 1 indicate that the point estimate may plausibly be too high or too low by the amount of the uncertainty factor. This is only one way to indicate approximate uncertainties, but is often useful for multiplicative models. Combining the quantified uncertainty factors using a central limit theorem (discussed in the following technical note) gives an estimated uncertainty factor of 18 to the point estimate of 1.84 cases per year. This provides the user with a sense of how many times larger or smaller than 1.84 the true but unknown rate of cases per year might be, based on the quantified uncertainties and point estimates in Table 2. In addition to these quantified uncertainties, as already mentioned, changing the 0.00595 point estimate of the fraction of severe cases (from Buzby et al., 1996), or changing the scoping assumption that only these severe *C. jejuni* cases warrant treatment with antibiotics and might receive clinical benefits from such treatment, could lead to proportional changes in the point estimate of risk.

Technical Note: Simplified multiplicative uncertainty factors. To enable quick approximate uncertainty analysis without Monte Carlo simulation, it is convenient to impose the artificial restriction that uncertainty about each parameter is approximated by a single multiplicative uncertainty factor. In other words, uncertainty about the point estimate x of an uncertain parameter X is expressed by an *uncertainty factor,* UF,

such that the true value of X is considered equally likely to be above or below its point-estimated value, x, and there is a subjective 95% probability that the true value of X lies between x/UF and x*UF. The interval [X/UF, X*UF] is interpreted as a subjective Bayesian confidence interval for X. Although this simplified approach to uncertainty assessment is not flexible enough to represent arbitrary beliefs (e.g., it is inappropriate for representing quantities with zero as a plausible value, or proportions for which X*UF > 1), it does allow uncertainty about each model parameter to be expressed, at least approximately, by a single number. Moreover, formulas for the human health benefits and risks from risk management interventions are often expressed as products of uncertain parameters, and may be expected to have approximately log-normal uncertainty distributions (Druzdzel, 1994). Uncertainty factors for components, say, u_1, u_2,..., u_n, of a product combine to yield the uncertainty factor for their product, via the formula:

$$\text{Uncertainty factor for product} = \exp\{2*[(.5*\ln(u_1))^2 + (.5*\ln(u_2))^2 + \ldots + (.5*\ln(u_n))^2]^{0.5}\}$$

(based on approximating the normal distribution on the log scale as a sum of normal distributions for the different components, each with an approximate 95% probability interval of 2 standard deviations.) For example, the uncertainty factors of 2, 5, and 10 in the top ("Exposure") section of Table 2 combine according to this formula to give a total uncertainty factor of 18 for their product. This uncertainty factor approach is used only to make uncertainty calculations more transparent. Monte Carlo uncertainty analysis is more flexible and general, but less easy to manually verify.

Technical Note: Monte Carlo uncertainty analysis. It is common practice to use Monte Carlo uncertainty analysis tools to assign probability distributions to factors such as those in Table 2 and to propagate these uncertainties through the calculations to obtain a probability distribution for the final calculated risk. A more general approach is useful for Bayesian Network and causal graph models in which the conditional probability distribution of the value of each node (representing a variable in the model) is determined by the values of the variables that point into it, perhaps with the values of some variables being observed as data. Then the conditional probability distribution for model variables can be calculated via well-developed exact computational algorithms (Zhang, 1998; Dechter, 1999). A simpler approximate method, now widely applied in health risk assessment modeling, is Monte Carlo simulation (Cheng and Drudzdel, 2000). If no Bayesian inference is required, then computer-aided risk assessment tools such as @RISK™, CrystalBall™ and Analytica™ can be used to sample values from the probability distributions of input nodes (nodes with only outward-directed arrows) and to propagate them forward through the deterministic formulas and conditional probability look-up tables (CPTs) stored at other nodes of the risk model to create approximate distributions for the values of output nodes (those with only inward-directed arrows). If Bayesian inference is to be used to condition on data while propagating input distributions to obtain output distributions, then specialized software such as the Bayesian Net Toolbox or WinBUGS can be used to perform the more computationally intensive stochastic sampling algorithms (typically, Gibbs Sampling and other Markov Chain Monte Carlo (MCMC) methods) required for accurate approximate inference in DAG models (Cheng and Drudzdel, 2000; Chang and Tien, 2002).

Table 2 has illustrated the calculations and an intermediate result for one risk management decision option: doing nothing, i.e., the *status quo* option. In other words, 1.84 cases per year is a preliminary upper-bound point estimate (ignoring uncertainty factors, health consequences, and preventable fractions, as explained above) of the number of severe macrolide-resistant *C. jejuni* cases that may be deprived of clinical benefits of macrolide treatment each year in the absence of intervention. To support rational decision-making, however, *it is essential to evaluate more than one option,* i.e., to inform the decision-maker about the consequences of alternative choices. (Policy makers do not always heed this principle. For example, an alternative approach to decision-making is to

invoke "situation-action" rules in which surveillance and monitoring data are used to trigger pre-specified intervention actions whenever certain conditions are detected, without calculating or comparing the probable human health consequences of alternatives. However, risk analysis generally focuses on and strives to support rational, i.e., consequence-driven, decision-making, and we will continue to make this assumption.)

The simplest alternative to the "do nothing" (i.e., *status quo*) option is to restrict or ban macrolide uses in chickens. The probable human health consequences of such an intervention, measured by the incremental number of illness-days caused or prevented, show the human health risks or benefits, respectively, of this option. Table 3 summarizes the main calculations and results. The detailed calculations and the symbols in the first column are explained later, in discussing Table 5; for the moment, the main point is just that a similar multiplicative framework to that used to calculate the risk estimate in Table 2 can be used to calculate the human benefits of the *status quo* (i.e., the human health risks from replacing it with a specified alternative). For the point estimates in Tables 2 and 3, continued use of macrolides is estimated to prevent about 6602.1 additional *C. jejuni* cases per year (39.3 of them severe) that would be created if macrolide use in chickens ceased, due to increases in airsacculitis-positive (AS+) chicken flocks. (Other potential sources of animal and human health benefits, such as decreased necrotic enteritis, are not considered in this example but could be quantified similarly.) This significantly outweighs the 1.84 severe macrolide-resistant cases per year estimated in Table 2 as being potentially preventable by changing the *status quo*.

Table 3: Example of a Top-Down Benefit Assessment for Macrolide Use in Chickens

HEALTH BENEFITS OF CONTINUED USE/ HEALTH RISKS CAUSED BY A BAN		
Variable	**Values and Uncertainty Factors (UF)**	**Data Sources**
Increase in fraction of chicken servings from airsacculitis positive (AS+) flocks if animal antibiotic use ceases, ΔF	0.005 assumed for macrolides as base case, assuming 50% effective substitution and no increase in prevalence rates	Historical data
Microbial load ratio factor	Estimated ratio of *C. jejuni* loads from AS+ compared to AS- birds = 10, UF = 10	Russell, 2003.
Dose-response ratio factor = ratio of risk-per-cfu for 10-fold greater doses compared to current doses	1 for linear no-threshold dose-response model; 0.3 for Beta-Poisson risk model; 13.93 for log-exponential model	Appendix B; Rosenquist et al., 2003; FDA, 2001
Incremental risk of campylobacteriosis per AS+ chicken serving, $(P^+ - P^-)$	Incremental risk = 10*(risk | AS-) – (risk | AS-) = 1.19E-4 per serving	Linear no-threshold model
Incremental *C. jejuni* cases per year caused by increased AS+ chicken servings	$\Delta F*MN(P^+ - P^-)$ = 0.005*38*292E6*1.19E-4 = **6602.1 additional cases/year.**	Product of above
Average health consequence per case	$sQ_s + (1 - s)Q_r$ = **6.128 days per case**	Marano, et al., 00
Additional illness-days per year	**40,458** = 6602.1*6.128	Product of above
Incremental *severe* cases per year	**39.3** severe cases/year = 0.00595*6602.1 additional cases caused per year	Buzby, et al., 1996

For completeness, Table 4 summarizes the potential human health risks and benefits from a withdrawal of macrolides in chickens, including the consequence component.

TABLE 4: Example of RRRT Risk-Benefit Assessment Calculations for Macrolides

Variable	Values and Uncertainty Factors (UF)	Data Sources
EXPOSURE: CURRENT RESISTANT CASES CAUSED BY ANIMAL ANTIBIOTIC USE		
Total current campylobacteriosis cases reported per 100,000 people per year	13.37 cases/100,000 in 2002 (for FoodNet surveillance sample)	CDC, 2003
Fraction of reported *C. jejuni* cases that are severe (treatment with antibiotic indicated)	0.00595	Buzby, et al., 1996
Average total severe cases per reported severe case	8 (Ranges from 2 for severe cases to 38 for mild cases; uncertainty factor = **5**)	Mead et al. (1999)
US population, N	292E6 = 2,920 x 100,000 people in US	US Census Bureau
Fraction of severe cases that are *C. jejuni*	0.99 (May be as low as 0.95)	CDC DBMD
Fraction of severe *C. jejuni* cases that are food-borne and caused by chicken products (including cross-contamination of other foods)	0.10, uncertainty factor = 3-**10**, subjective estimate based on competing risk, genetic, epidemiological, historical data and bounds	Appendix B; Stern and Robach, 2003
Fraction of chicken-caused severe cases that are antibiotic-resistant, (1 – s)	0.01 for erythromycin resistance, UF = **2**	CDC, 2000
Preventable resistance fraction = Fraction of resistant chicken-caused severe cases that would become susceptible to analogous human antibiotics if animal antibiotic use ceased, p.	≤ 1. (True value could be 0 based on temporal trend evidence before and after withdrawals. See discussion in text.) Point estimate: 0.3, UF = 3	Hayes and Jensen (2003); Gaudreau and Gilbert, 2003, 1998.
Resistant severe *C. jejuni* cases per year caused by chicken products = (P^-)*(MN)*(1- s) = product of above factors	**1.84** cases/yr. for macrolides = (13.37E-5) * 0.00595*8*292E6*0.99* 0.10*0.01; UF = 18 (from component UFs of 5, 10, 2)	Product of above.
CONSEQUENCE: PREVENTABLE CURRENT HUMAN HEALTH CONSEQUENCES OF RESISTANCE		
Fraction of resistant cases that are given an antibiotic to which they are resistant, r	~0.5, UF = 2	FDA-CVM, 2001
Fraction of *in vitro* "resistant" chicken-caused severe cases that do not achieve normal clinical benefit from treatment with resisted antibiotic, f	≤ 1	Upper bound
Human health harm per severe resistant case treated with antibiotic that would be prevented if it were replaced with a susceptible case = $(Q_r - Q_s)$ days of illness prevented	2 illness-days (assumed baseline value, starting point for sensitivity analysis) for severe cases; 0 days for non-severe and untreated cases	Ang and Nacham, 2003. See discussion in text.
Population risk = Preventable resistant cases (and illness-days) of severe *C. jejuni* per year expected in US $p[(1- s)*(P^-)*(MN)]rf(Q_r - Q_s)$	≤ **1.84** illness-days/yr. for macrolides;	Product of above factors
HEALTH BENEFITS OF CURRENT USE: INCREMENTAL HEALTH RISKS CAUSED BY A BAN		
Increase in fraction of chicken servings from AS+ flocks if animal antibiotic use ceases, ΔF	0.005 assumed for macrolides as base case	Assumes 50% effective substitution, no increase in prevalence rates
Microbial load ratio factor	Estimated ratio of *C. jejuni* loads from AS+ compared to AS-birds = 10, UF = 10	Russell, 2003.
Dose-response ratio factor = ratio of risk-per-cfu for 10-fold greater doses compared to current doses	1 for linear no-threshold dose-response model; 0.3 for Beta-Poisson risk model; 13.93 for log-exponential model	Appendix B; Rosenquist et al., 2003; FDA, 2001
Incremental risk of campylobacteriosis per AS+ chicken serving, $(P^+ - P^-)$	Incremental risk = 10*(risk | AS-) – (risk | AS-) = 1.19E-4 per serving	Assumes linear no-threshold model
Incremental *C. jejuni* cases per year caused by increased AS+ chicken servings	$\Delta F*MN(P^+ - P^-)$ = 0.005*38*292E6*1.19E-4 = **6602.1 additional cases/year.**	Product of above
Average health consequence per case	$sQ_s + (1 - s)Q_r$ = **6.128 days per case**	Marano, et al., 2000
Additional illness-days per year	**40,458** = 6602.1*6.128	Product of above
Incremental *severe* cases per year	**39.3** severe cases/year = 0.00595*6602.1	Buzby, et al., 1996

OTHER CONSIDERATIONS AND EXTENSIONS OF RRRT CALCULATIONS

In addition to its direct effects on resistance levels and pathogen loads reaching consumers, withdrawal of an animal antibiotic may have important indirect effects that depend in part on the decisions and behaviors of human stakeholders as they adapt to the ban and/or that are transmitted via other causal pathways than those addressed in the model. Examples of such additional considerations, with brief comments, are as follows.

- *Antibiotic substitutions and synergies.* Following the ban on antibiotics used as growth promoters in food animals in Europe, therapeutic use of other animal antibiotics to treat animal diseases increased significantly (Casewell et al., 2003). By analogy, withdrawing macrolide use in the US might cause an increase in airsacculitis-positive (AS+) flocks that could be treated by veterinary prescriptions of enrofloxacin. Conversely, withdrawing enrofloxacin might be compensated for by increasing use of macrolides to prevent infections leading to AS+ flocks. If both macrolides and fluoroquinolones were withdrawn, however, the increase in AS+ flocks might increase far more than if either one alone is withdrawn.
- *Other animal bacterial diseases.* Macrolides and streptogramins (virginiamycin) are effective against necrotic enteritis (NE) caused by *Clostridium perfringens* (e.g., Brennan et al., 2001; Vissiennon et al., 2000). Following the 1999 ban on these and other growth promoters in Europe, NE rates in some countries increased sharply before settling to new, higher levels with increased use of therapeutic drugs (e.g., Lovland and Kaldhusdal, 2001; Madsen and Pederson, 2000). If human health risks from NE+ flocks are comparable to those from AS+ flocks, then the human health benefits from continued use of macrolides and virginiamycin to control NE may be significant.
- *Other foodborne human pathogens.* This example RRRT assessment has focused on *C. jejuni*. Although *C. coli* cases are only a small percentage of total campylobacteriosis cases, they have much higher resistance rates – 22.5% against erythromycin, compared to 0.5% for *C. jejuni*, according to Fedorka-Cray et al, 2001. If resistance rates in *C. coli* are about 45 times as great as for *C. jejuni* and *C. coli* constitute a few percent of the total cases, then the human health benefits from withdrawing macrolides could be about double those estimated in Table 4 for *C. jejuni*. Conversely, Russell (2003) reported significantly greater average loads (although not on every replicate) of *Salmonella* and other pathogens, as well as *C. jejuni*, in processed carcasses from AS+ compared to AS- flocks. The human health impacts from these other pathogens might significantly increase the estimated human health benefits from continued use of animal antibiotics. The RRRT calculations could help to quantify these additional effects.
- *Co-selection and commensals.* Macrolide use in chickens may co-select *E. faecium* that are resistant to streptogramins (SREFs), although the genes responsible for resistance to streptogramin A are rarely found in animal isolates. Since glycopeptides and linezolid are not used in food animals in the US, there is no potential for co-selection with genes that confer resistance to these two human-use only antibiotics. A risk assessment for virginiamycin (Cox and Popken, 2004, discussed below) indicates that these risks are quantitatively very small (fewer than 1 expected excess mortality in 20 years). The contribution from continued use of macrolides is therefore predicted to be less than the rounding errors in Tables 3 and 4.
- *Reduced need to treat human patients with antibiotics.* The bottom of Table 4 estimates that a ban or risk management intervention restricting continued use of macrolides could increase total campylobacteriosis by about 6602 cases per year. Perhaps about 10% of these cases might seek and receive treatment with ciprofloxacin or macrolide antibiotics, almost entirely as empiric treatments. Preventing these cases would remove these human antibiotic prescriptions, potentially reducing selection pressure for resistance in human pathogens and commensals.

- *Changes in prescription practices.* As physicians and scientists become more concerned about avoiding over-prescriptions of antibiotics with doubtful clinical benefits, more rapid and accurate diagnostic and resistance-screening tests may continue to be developed (e.g., Endtz et al., 2000). These could reduce the prescription rate for resisted antibiotics in humans, and hence reduce the potential benefits of withdrawing animal antibiotic use.
- *Opportunistic infections and patient practices.* If people treated with antibiotics for other (non-campylobacteriosis) reasons are thereby made significantly more vulnerable to infection by antibiotic-resistant Campylobacter ingested in chicken or other foods, then the benefits of a ban might be understated in the analysis in Tables 3 and 4 (essentially because P^{-} would be greater than estimated for resistant campylobacter-contaminated servings). As explained in Appendix B, empirical evidence (e.g., Friedman et al (2000), Smith et al (1999) data) does not show a positive association between chicken consumption and risk of resistant campylobacteriosis, but, the possibility cannot be ruled out. Increasing awareness by patients and at-risk individuals of the need for care in food preparation, cooking, and handling would tend to attenuate any benefit from this hypothesized source.
- *Emergence of resistance.* A common concern is that the resistance fraction $(1 - s)$ may drift up over time unless animal antibiotic use is curbed (FAAIR, 2002). However, biomathematical modeling suggests that, at least for antibiotics such as virginiamycin and macrolides that have been used for several decades in food animals without leading to high levels of resistance in people, an outbreak of high resistance in the future from this source is very unlikely (Cox and Popken, 2004b.)
- *Timing:* For simplicity, and to be conservative (i.e., maximizing the estimated risk of continued use of the animal antibiotic) the *timing* of human health impacts of a ban has so far been ignored: only the new levels that will eventually be reached have been considered. Evidence from Europe suggests that the hypothesized health benefits to human patients from banning animal antibiotics may take longer than 5 years to materialize (Heuer et al., 2002; Borgen et al., 2000, Iversen et al., 2002), while adverse impacts on increased animal pathogen loads (e.g., Madsen and Pederson, 2000) and possibly on human health (*Eurosurveillance, 2000*) may be much more immediate. If so, then modeling the timing of impacts might further increase the benefit-to-risk ratio for continued use of animal antibiotics in this example.

In summary, while the analysis in this example has focused on *C. jejuni* transmitted via chicken servings, other important considerations may tend to strengthen the conclusion that human health risks from withdrawing or restricting macrolide use in chickens could significantly outweigh potential human health benefits. Such additional comparisons and information can be included in an expanded quantitative human health risk analysis, but if their main effect is to strengthen the already strong comparison of options in Table 4, then the additional resources and effort required to quantify them further may not be worthwhile, i.e., better information on these points may not lead to any change in the relative evaluation of decision options. In this case, the additional information in further quantitative risk assessment would have no incremental value for risk management decision-making. A key prescriptive principle of value-of-information analysis is not to pay for information that does not have the potential to change the risk management decision.

CRITERIA FOR A SUCCESSFUL RISK ANALYSIS

A successful risk analysis does the following:

- *Scope the analysis to support decisions* by estimating the causal relation between decisions, exposures, and their probable total human health consequences. To guide rational regulatory decision-making, traditional quantitative risk analysis seeks to quantify the causal relation between regulatory actions that might be taken and their total probable human health consequences.

- *Evaluate proposed solutions, not problems.* The risk analysis should yield qualitative and/or quantitative evaluations of proposed risk management actions. A successful risk analysis shows the estimated frequencies and magnitudes (and uncertainties) of human heath consequences caused by different proposed risk management decisions. It is important to identify an adequate range of risk management options to assure that dominant alternatives are not overlooked.

- *Evaluate total human health impacts.* Total health consequences are found by summing the impacts of proposed actions on human exposures to microbial loads of bacterial species (both resistant and susceptible) over all relevant pathways that contribute significantly to the outcome (e.g., different food animal species, drinking water, home-cooked meals, restaurant dining, etc.) Applying a qualitative or quantitative exposure-response model to the changed exposures for different decisions then yields the estimated risks associated with them.

- *Communicate clearly and enable effective participation.* A well-conducted risk analysis enables its recipients to participate more effectively in risk management deliberations and to communicate questions and concerns more clearly and concisely than would otherwise be possible. It does so by providing the relevant information needed to determine the probable consequences of proposed actions and by showing how sensitive these predicted consequences are to specific uncertainties and assumptions in the analysis.

Bailar and Travers (2002) suggested additional pragmatic criteria, including:

- Reduced demand on resources
- Common format
- Reduced demands for data
- Easy comprehension by non-experts; and
- Ready adaptation.

They recommend a multiplicative model, similar in concept to the RRRT approach in Table 4, to meet these criteria. They state that such a model should estimate the annual number of symptomatic infections by the organism of interest in a specific population; the fraction of those occurrences in which the bacterial strain was clinically resistant to the

antimicrobial or class of antimicrobials under study; the annual number of occurrences in which infection by a resistant strain led to the specific adverse health outcome(s) under study; and the fraction of the adverse outcomes in which the antimicrobial resistance was a result of the farm use or category of uses under study. These four factors are a subset of those considered in Table 4, which also addresses the preventable fraction of such cases, i.e, the fraction that could be prevented (or caused) by a change in animal antibiotic use.

The following sections address the components of risk assessment, risk management, and risk communication more fully, with emphasis on risk assessment.

RISK ASSESSMENT

DEFINITION OF RISK ASSESSMENT

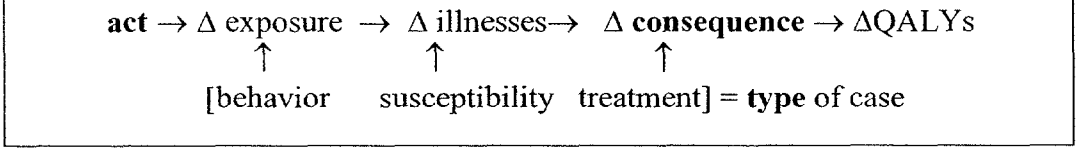
Risk assessment is defined by the *Codex Alimentarius* Commission as "A scientifically based process consisting of the following steps: (i) Hazard identification, (ii) Hazard characterization, (iii) Exposure assessment, and (iv) Risk characterization." The following sections discuss each step in this process. Throughout the discussion, "scientifically based" is taken to mean: "Based on specifically identified, independently verifiable data sources and on explicitly stated, empirically testable hypotheses, models, and calculation formulas or algorithms."

ANIMAL ANTIMICROBIAL RISK ASSESSMENT: CONCEPTUAL FRAMEWORK

To support effective risk management decisions, human health risk assessments must characterize known or suspected potential causal relations between proposed risk management actions and their probable human health consequences. The actions typically affect exposures to sources of risk (i.e., "hazards"), while the consequences typically include changes in frequency or severity of resulting illnesses or deaths in the affected population and perhaps in various at-risk sub-populations. Impacts of changes in animal drug use can potentially be transmitted to humans by several causal paths, such as changes in exposures to microbial loads of both susceptible and resistant strains of bacteria in food commodities, and perhaps transfer of resistance determinants to humans via these or other bacteria. The medical consequences of changes in exposures to microbial hazards will depend on the resulting changes in illness rates, on patterns of resistance to human drugs among cases of food-borne illness, and on treatment and prescription patterns for patients receiving human antibiotics. Hazard identification consists of identifying the various causal paths that lead from risk management actions to their human health consequences.

Figure 1 outlines a causal graph template (Shipley, 2000) for assessing risks to humans from changes in animal drug uses. In this template, risk management *actions* that change current practices or activities such as animal drug use can thereby change *exposures* of individuals to potentially harmful agents (the hazards, typically one or more bacterial strains). Changes in exposures, in turn, change expected *illness rates* and hence adverse *health consequences* (e.g., illness-days or early deaths per capita-year) in susceptible members of the exposed population. If desired, different human health consequences can be aggregated into a single summary measure such as quality-adjusted life-years (QALYs) (Hazen, 2003), although this is optional. Effects of changes in animal drug use on QALYs lost per year in the population may be mediated by behaviors (e.g., kitchen hygiene, cooking, and care-seeking behaviors), individual attributes (e.g., immune status, age, sex, and other covariates that affect susceptibility to infections), and physician prescription practices. These covariates may also influence each other (indicated by the brackets [] around them in Figure 1.) For example, an AIDS patient may have food consumption and preparation behaviors and receive medical care and prescriptions different from those of a non-AIDS patient. Risk management options (acts) are sought that decrease adverse health consequences, taking into account the distribution of covariates in the population.

FIGURE 1: A Causal Graph for Health Risk Analysis



Technical Note: Bayesian Network risk model. A useful mathematical and statistical framework for Figure 1 interprets it as a Bayesian belief network (BN) or causal graph model (Chang and Tian, 2002). Each variable with inward-pointing arrows is interpreted as a random variable with a conditional probability distribution that depends only on the values of the variables that point into it. The essence of the forward Monte Carlo approach to modeling and evaluating uncertain risks in this framework is to sample successively from the conditional distribution of each variable, given the values of its predecessors (*ibid*). Important microbiological processes, such as cross-contamination during processing of animal carcasses, are represented in such models only implicitly, by conditional probability distributions of microbial loads on outgoing (processed) carcasses, given the microbial loads on incoming carcasses. Algorithms exist to identify and validate possible causal graph structures from data (e.g., Tsamardinos et al., 2003) but are not yet routinely applied in risk assessment.

For practical purposes, each choice of a risk management **act** in Figure 1 will generate an approximately Poisson-distributed number of incremental illness cases ("responses") caused or prevented each year in each severity class (e.g., mild, moderate, severe, fatal) in the population (and in each sub-population, if there are several). The expected health consequences of this change can be calculated from the following three components models, common to most risk assessments:

- An **exposure model** (the "**act → Δexposure**" link in Figure 1) that quantifies the average number of contaminated servings ingested per year, for population risk; or average contaminated servings ingested per capita-year, for individual risks. "Contaminated" here means carrying enough pathogenic bacteria (possibly just one) to pose an elevated risk of food-borne illness to susceptible consumers. The number of contaminated servings ingested per year is also approximately Poisson-distributed, and so is fully characterized by its mean. The exposure model may depend on a consumer's "type", i.e., on individual covariates such as food purchasing, preparation and consumption variables that affect exposures.
- A **dose-response or exposure-response model** (the "**Δexposure → Δillnesses**" link in Figure 1) that quantifies the probability of illness or expected number of cases of a given severity (for infectious illnesses) per contaminated serving ingested. In general, this may also depend on the consumer's "type", i.e., on the combination of covariate values that affect risk for that individual, as well as on the dose ingested.
- A **health consequence** model (the "**Δillnesses → Δconsequence**" link in Figure 1) quantifying probabilities of different health outcomes (e.g., survival vs. fatality, QALYs lost) from each case. These outcome probabilities may depend on physician prescription practices for different types of cases.

These three sub-models determine the expected annual number of human illness cases in each severity class and the expected QALYs lost per year, for each choice of **act**.

In general, there may be several distinct bacterial strains, food animals and commodities, and at-risk sub-populations (perhaps including groups receiving different medical treatments) that are to be included within the scope of the risk assessment, since the goal of risk assessment is usually to quantify the *total* human health impact of interventions such as changes in animal drug use. In this case, summing health impacts over all combinations (each corresponding to an instance of Figure 1) gives the total probable change in human health consequences for the act.

The framework in Figure 1 can be implemented with more or less sophistication. Perhaps the simplest useful approach is to estimate the following three factors for each risk management act and path being evaluated:

- ***Exposure factor*** = average contaminated servings ingested per capita-year
- ***Dose-response* factor** = expected cases of illness per contaminated serving ingested
- ***Health consequence factor*** = expected QALYs lost (or illness-days, etc.) per case of illness. (Alternatively, a vector of expected numbers of different health outcomes, e.g., mild, moderate, severe, and fatal illnesses per case, can be estimated.)

If these factors are multiplied by each other and by the number of people affected for each causal path (i.e., each bacteria-food-human sub-population combination of interest) for a risk management action, and if the resulting products are summed over all causal paths, then the sum provides an estimate of the total human health impact per year for that action. A more refined calculation can be made by considering how the factors are likely to change over time and then summing over time periods (perhaps with discounting). A simpler expedient is to assess and compare the steady-state equilibrium annual risks for different risk management scenarios after all transients have settled down.

At the other end of the sophistication spectrum, instances of Figure 1 can be assessed and applied to risk estimation problems using conditional probability calculation techniques developed for Bayesian Networks and causal graphs (Chang and Tian, 2002; Tsamardinos et al., 2003; Shipley, 2000). In this case, hazard identification consists of verifying that the causal path in Figure 1 leading from acts that change animal antibiotic uses to resulting changes in exposures, responses, and health consequences is consistent with available data. The remaining steps in the risk assessment process can be interpreted as quantifying and applying the Bayesian Network model. Within the Bayesian Network framework, the simple conceptual model of multiplying exposure, dose-response, and consequence factors together, as in Figure 1 and Table 4, generalizes to allowing arbitrary probability distributions for inputs and conditional probability relations or functions at the nodes to be combined by Monte Carlo uncertainty analysis to derive the joint probability distributions of outputs. The Bayesian Network modeling perspective is potentially very useful, but is not yet widely adopted in animal antibiotic risk analysis.

BACKGROUND: SOME PREVIOUS RISK RATING APPROACHES

Motivated in part by concerns that quantitative risk assessment of human health risks from animal antibiotic use (AAU) might prove to be overly burdensome to implement, insufficiently credible, and/or require data that are not readily available or

assumptions of doubtful validity to bridge data gaps, several regulatory risk analysis groups have proposed qualitative rating approaches designed to avoid these pitfalls. For example, a three-component risk rating with components of "Hazard", "Exposure" and "Impact" has been developed in Australia. Risk is profiled with the help of the following 3 x 4 matrix:

Australia National Registration Authority Veterinary Requirements Series Part 10

Factor	Definition	Negligible	Low	Medium	High
Hazard = source of risk	**Antibiotic resistant microorganisms** or their resistance plasmids (that have the potential to transfer to humans) within an animal species, arising from the use of an antibiotic in an animal species				
Exposure	**Amount and frequency of exposure** of susceptible humans to antibiotic-resistant microorganisms (or their plasmids) from animal sources				
Impact	The evaluation of infections (caused by antibiotic-resistant pathogens of animal origin) in susceptible humans. Considers: a) Perceived or known **clinical importance** of the class of antibiotics to humans; b) **Dose response** assessment of relationship between frequency and magnitude of exposure of humans (dose) to antibiotic – resistant food-borne microorganisms and severity and/or frequency of the impact (response); including an estimate of the critical threshold of exposure required to cause infection in susceptible humans. c) Antibiotic-resistant **disease severity, morbidity, mortality**. d) **Expected numbers** of **infections and deaths**. e) The impact on human health and **quality of life** including the range of the **susceptible humans** expected to be affected. Probability of antibiotic-resistant infection development in susceptible humans (N, L, M, H)				

Source: Adapted from http://www.apvma.gov.au/guidelines/vetguideline10.pdf.

Separate risk summaries may be required for different bacterial species. The assessment also includes:

- **Uncertainty** of data used in risk assessment, including both a) Uncertainty due to inherent variability and measurement error; and b) Uncertainty due to lack of information or understanding.
- **Benefits** of use of antibiotic in Australian animal health; and
- **"Risk"**, characterized as *"Probability of disease due to infection in susceptible humans after exposure of humans to antibiotic-resistant microorganisms (and genetic material) of animal origin* ***and*** *the severity of the impact of exposure on susceptible humans".*

This framework contains many potentially useful ideas, including consideration of AAU benefits and uncertainties about AAU risks as part of the assessment; estimation of expected illnesses and deaths; distinction among illnesses of different severities; and identification of (perhaps multiple) susceptible sub-populations and multiple bacterial species if required to adequately characterize risk. However, the impact category contains items (e.g., dose-response relation and clinical importance of human antibiotics) that might be redundant once the expected number and severity of additional morbidities and mortalities caused by a change in AAU are known. In other words, dose-response and

clinical importance can be considered means to an end: predicting the change in human health impacts from a proposed change in AAU. Once the human health impacts are known, the factors used in calculating them are no longer needed to characterize risk. Moreover, the conceptualization of risk as "Probability of disease due to infection in susceptible humans after exposure of humans to antibiotic-resistant microorganisms (and genetic material) of animal origin **and** the severity of the impact of exposure on susceptible humans" may not be entirely satisfactory. For example, consider an extreme hypothetical case in which resistant strains have no ability to cause illness (zero virulence) and all illnesses are caused by susceptible strains – but susceptible and resistant strains typically occur together in infected patients. It is not clear that the risk concept defined here would attribute zero (or negligible) risk to the resistant microorganisms in such a case, even if they have no adverse effect on human health. Thus, in seeking to create a new risk rating system, it may be necessary to refine the concept of risk, even while using many of the ideas in the above framework.

Canada has a somewhat similar qualitative rating system for risk analysis of plant pests, including bacteria, again using H = high, M = medium, L = low, N = negligible for risk, its components, and its impacts. Key aspects of this rating system are as follows:

> "Probability may be estimated under two broad scenarios: 1) In the baseline (or uncontrolled) scenario, probability is estimated under natural or status quo conditions. It is estimated in the absence of artificial or un-natural spread of the pest, and in the absence of "new" risk management actions being taken, beyond the status quo. 2) Alternatively, it may be appropriate to estimate probabilities under proposed risk management (or controlled) scenarios. The probability of the pest's establishment is considered and succinctly described under three headings, including: a) the potential for its entry into the area through various pathways, b) the extent to which suitable hosts and habitat are available, and c) its potential for spread from the initial point of entry. ... The final three headings of the rating system are concerned with impact or consequence of the pest, and include: **a)** the range of hosts, **b)** direct and indirect economic consequences, and **c)** general environmental impact. ...The above probability and impact estimates are summarized in an overall risk rating of **negligible, low, medium** or **high**. Subsequently, the sources and magnitude of uncertainty of the estimation of risk are summarized in a description of the assumptions used in the estimation, a discussion of the nature and quality of the data, and a discussion of the supporting and conflicting evidence. Finally, a concise statement is included, noting whether the pest(s) or commodity represent a health hazard to humans or animals."
>
> *Source:* http://www.gov.on.ca/OMAFRA/english/research/risk/frameworks/as4c.html

An important aspect of this system is that it considers the change in estimated probabilities of risk components if different risk actions are taken. This concept – using risk rating systems to link proposed risk management actions to their probable consequences, defined as changes in the probabilities (or of statistical frequencies in affected populations) of the outcomes of interest – can be applied to many settings other than the plant pest context. In particular, it suggests that the human health risk of a proposed change in AAU, such as introduction of a new product or withdrawal of an existing one, should be assessed by considering how human health impacts are likely to change if the proposed action is taken. This emphasis on the human health consequences of risk management decisions is

consonant with many recommendations that risk analyses should be decision-focused and provide information useful for assessing risk management decision options.

The Canadian approach to qualitative risk rating has been extended to food safety risks. This framework is described as follows:

Ontario Ministry of Agriculture and Food Risk Assessment Framework

"Risk assessors are responsible for risk characterization. The risk is characterized by estimating in qualitative or quantitative terms, the probability of and the magnitude of the impact (or consequence) of the adverse effects of the disease. The risk is further characterized by noting the attendant uncertainty of the estimates, given the available data.

When reliable quantitative data is available, assessors use quantitative multiplicative mathematical models to estimate risk. Often, the desired quantitative data are not available. In such cases a more qualitative approach is used. In either case, quantitative and qualitative assessments are summarized using a rating system to help categorize risks. The final rating assigned to a given hazard / commodity situation, is derived from six sub-ratings, each rated as negligible, low, medium or high.

The first three sub-ratings are concerned with the probability of a human health impact being realized. This is influenced by several factors including the exposure characteristics of the situation. The final three sub-ratings are concerned with the impact of the disease, which is influenced by several factors including dose-response characteristics. This scoring system is used to help categorize risks in terms of their general importance. It is not used to rank individual risks in numerical sequence, but does attempt to place them in broad categories of negligible, low, medium or high risk. ...

The probability of exposure is considered under three headings: **a)** The **probability of contamination** of food along the food chain, by disease agents. It may require consideration of pathways for contamination of source animals or crops and of the product during processing, storage, distribution and preparation, **b)** The **probability of significant exposure** of susceptible human hosts to a dose sufficient to cause disease, **c)** The **potential for broad distribution and/or secondary spread** of the disease....

The impact (or consequence) of a hazard is the second component of risk. Currently, only negative human health impacts need to be described in a food safety risk assessment for the purposes of trade. Therefore, the emphasis of OMAF impact-assessment is on impact to human health. Disease impact is described in terms of its severity and frequency of debilitation, and its impact on quality of life. The OMAF model also briefly describes economic and environmental impacts of food safety hazards.

Risk characterization and estimation are summarized in a concise statement noting the probability and impact of disease. The sources and magnitude of uncertainty of the estimation of risk, are summarized in a description of the nature and quality of the data, and a discussion of the supporting and conflicting evidence. If possible, the results of sensitivity and importance analysis of a quantitative mathematical probability model are summarized. Finally, an overall risk rating of **negligible**, **low**, **medium** or **high** is assigned.

Source: http://www.gov.on.ca/OMAFRA/english/research/risk/frameworks/as2c.html#1.0%20Risk%20Assessment1841

This framework retains the feature of comparing probabilities of consequences with and without different risk management interventions. A key feature is its use of a *multiplicative* approach to aggregate the components of the risk rating when adequate data are available.

Thus, a microbial hazard that creates zero human exposure or for which exposure has zero human health impact could have a risk rating of zero even if other factors were very large.

A more quantitative rating approach is the Brenner Scheme proposed in the U.K. for genetically modified microorganisms. (Since the new GM 2000 regulations, the Brenner Scheme is not used as the sole basis to classify risks for GMOs, but we are interested only in the aspects of it that might be useful for devising improved AAU risk rating approaches.) In this scheme, each of the following three factors is assigned an order-of-magnitude weight (e.g., 10^{-3}, 10^{-6}, or 10^{-9}):

- ACCESS = probability that the GMO or DNA contained within it will be able to enter the human body and survive there.
- EXPRESSION = measure of the anticipated or known level of expression of the inserted DNA
- DAMAGE = measure of the likelihood of harm being caused to a person by exposure to a GMO independently of access and expression

These three risk factors are then multiplied together to give an overall risk factor, which is then used to look up the containment level required for the experiments. (*Source:* http://www.biology.ed.ac.uk/sbs/healthsafety.htm.)

The Brenner approach to GMOs, like the Canadian framework for food safety assessments, emphasizes a *multiplicative* approach for aggregating components of risk into an overall risk factor or risk score. Unlike the previous frameworks, however, it uses order-of-magnitude estimates of the components, rather than qualitative labels (H, M, L, N). The final risk rating is also used to look up the required risk management (containment) approach. Key ideas include using rough numerical estimates of risk factors or risk components; multiplying the results to get a rough quantitative estimate of overall risk; and mapping this rough quantitative risk estimate into a risk management decision category.

Lessons from Previous Approaches

Comparing the risk rating systems above suggests valuable components to include in any (qualitative or quantitative) risk rating system for animal antimicrobials. These are listed and discussed next, with particular reference to the US FDA's Center for Veterinary Medicine (CVM's) Guidance #152 document, which sets forth an approach to qualitative risk rating similar in some respects to those mentioned above. Some aspects of risk rating systems that might be useful to consider in designing future ones are as follows.

1. **Change in the *frequency* of adverse human health impacts** for different risk management decisions. The current definition of risk in the US CVM's Guidance #152 is *probability of "human illness that is caused by a specified antimicrobial-resistant bacteria (sic), is attributable to a specified animal-derived food commodity, and is treated with the human antimicrobial drug of interest."* This definition makes no reference to the effects of any risk management decisions that the risk assessment might help to support. In addition, the *probability* that a human illness is caused by a specified antimicrobial-resistant bacterium (e.g., for at least some member of a

population under at least some conditions) is less important than the *frequency* with which the illness occurs. A bacterium that is known with certainty to cause a very infrequent illness (probability = 1) may be of less concern than one that is only suspected of causing a very frequent or serious illness.

Whether used in quantitative or qualitative risk ratings systems, the conceptual units of frequency are *expected number of illnesses per year* (in an identified exposed population), for population risks; and *expected number of illnesses per capita-year* for individual risks.

Technical Note: Population heterogeneity. Frequencies should ideally be estimated for relatively homogeneous sub-populations, i.e., sub-populations whose members have approximately equal risks; otherwise inter-individual heterogeneity in risks must be addressed. Statistical techniques such as classification tree analysis and finite mixture distribution modeling can help to identify homogeneous sub-populations and to estimate frequencies for them from case-control, cohort, and longitudinal survey data.

2. ***Severity* of adverse human health impacts** from different risk management decisions (e.g., proposed changes in AAU). Currently, the "Consequence" rating in Guidance #152 refers to the "importance" of human drugs, but not to the adverse human health *consequences* caused by AAUs. As suggested by the Impact portion of the Australia table, the severity of human health impacts from preventable illnesses should be a key component of the risk assessment. The conceptual units for severity are: *expected adverse impacts per illness* (e.g., mortalities, morbidities, illness-days, life-years lost, etc.), perhaps with morbidities further broken down by severity class (e.g., mild, moderate, severe) and with mortalities further classified by age group or number of life-years lost. *Quality-adjusted life-years (QALYs)* lost may also be used if the required assumptions (Hazen, 2003) are accepted and it is desired to aggregate diverse health impact metrics into a single summary measure. (As in the case of frequency, severity of health impacts should also ideally be assessed for multiple sub-populations, e.g., based on age, immune status, etc., if impact severity distributions differ significantly among them. Classification tree analysis and other modern statistical methods can help to identify relevant subpopulations from data if available.)

The motivation for considering severity of human health impacts in rating risks is illustrated by the following example. Suppose that "probability of human illness caused by a specified resistant bacteria, attributable to a specified animal-derived food commodity, and treated with the human antimicrobial drug of interest" = 1, but that treatment with the human antimicrobial drug of interest is completely effective clinically (i.e., resistance makes no difference to clinical outcome). This situation should presumably be rated as less bad than one in which the probability is less than 1 but the impact is treatment failure and death due to resistance. To assure that the second situation is rated as worse, human health impacts must be considered.

3. ***Causality* of adverse human health impacts by proposed *changes*** in AAU. As suggested by the Canadian approach, it is useful to be able to assess the change in expected adverse human health consequences caused specifically by proposed risk

management interventions. Consider again the Guidance #152 definition of hazard as: *human illness caused by a specified resistant bacteria, attributable to a specified animal-derived food commodity, and treated with the human antimicrobial drug of interest* = 1. If the presence of the resistant bacteria in the specified food commodity is not actually *caused* by an AAU, then it should not be considered part of the risk from the AAU. To take an extreme hypothetical example for clarity, suppose that a proposed ban on the associated AAU would have the net effect of causing *more* resistant bacteria to be ingested per capita-year (e.g., because the AAU is not the sole source of resistance and because withdrawing it would amplify the total microbial loads, including the resistant portion, reaching consumers.) This information might be important for risk management decision-making, but would not necessarily be apparent if only the risk of the italicized hazard above is assessed.

4. ***Uncertainty*** about the changes in frequency and severity of adverse human health effects caused by a proposed change in AAU or other proposed risk management intervention. For example, what overall rating should be assigned to a situation that has a 50% chance of an "L" risk rating, a 30% chance of an "M" rating and a 20% chance of an "H" rating, depending on how scientific uncertainties are resolved? In the Canadian system, uncertainty is summarized along with risk characterization information before a final overall risk rating is applied. In the Brenner system, uncertainty about the component ratings is indicated by order-of-magnitude estimates and these uncertain estimates are then used to identify risk management responses. In the current CVM #152 Guidance, some uncertainty can perhaps be subsumed into the qualitative rating labels. In all of these systems, more explicit guidance on how to treat uncertainties in component ratings would be useful.

5. ***Cumulative risk assessment,*** i.e., total risk summed over the multiple pathways by which changes in AAU propagate to cause resulting changes in exposures to microbial loads and consequent adverse human health effects. These pathways may include multiple bacterial species and/or multiple drugs to which co-resistance or cross-resistance may be increased by the AAU (or proposed change in AAU) whose risks are being assessed; multiple food products; and perhaps multiple human sub-populations affected. They may also include susceptible as well as resistant strains of bacteria if both are affected by the proposed change. The goal is to consider *all* major pathways by which the proposed change or intervention leads to significant changes in human health impacts, so that the *total* human health impact can be considered.

6. ***Potential benefits*** to humans and animals from AAU or AAU changes. To inform rational risk management, the change in human health benefits, if any, from a proposed change in AAU must be assessed as well as changes in human health risks. Animal health benefits can also be listed separately in the overall assessment of likely impacts of proposed risk management interventions, as in the Australian system.

7. ***Necessary and sufficient information; flexible estimation procedures and information requirements.*** Guidance #152, like the other systems considered, lists

many potentially relevant and informative data elements to be considered in the rating process. Exactly how these data elements should be assembled to build up a coherent account of the overall human health risk caused by a propose change in AAU is less clearly specified. It is therefore possible that several overlapping or partly redundant pieces of information that address essentially the same bottom-line concern (e.g., exposure, response probability, etc.) might be considered while leaving unaddressed other key information (e.g., on the human health impacts specifically caused by resistance-related treatment failures) needed for decision-makers to understand how changes in AAU will affect human health risks.

8. ***Multiplicative aggregation.*** Guidance #152 is currently based on a look-up table that relates component ratings to overall risk rating in a pattern that could be interpreted as additive. By contrast, a multiplicative aggregation approach would allow the overall risk to be zero (or rated N = negligible) if any of its key components of exposure, exposure-response probability, or consequences is rated 0 (or N).

   As an example, suppose that the following ratios can be either rated (e.g., using an H, M, L, N scale) or estimated, perhaps to the nearest order of magnitude:

   - *Exposure Factor* = (Δ Exposure/Δ AAU) = (change in contaminated meals ingested per year) per (incremental animal treated with or exposed to the animal drug of concern). (Here, "contaminated" means contaminated with an infectious dose, i.e., one that is large enough to be cause illness in a susceptible exposed individual.)
   - *Exposure-Response Factor* = (Δ Illnesses / Δ Exposure) = (expected number of additional illnesses per year) per (contaminated meal ingested)
   - *Consequence factor* = (Δ Human Health Impacts /Δ Illnesses) = (expected number of adverse health consequences) per (illness case resulting from ingestion of a contaminated meal). If multiple impacts are considered, then separate consequence factors can be estimated for the different types of impacts (e.g., illness-days by severity category, mortalities, QALYs lost, etc.)

Then for a given change ΔAAU in animal antibiotic use that leads to a release of bacteria from the farm and into the food commodity stream (perhaps corresponding roughly to the *Release Factor*), the corresponding human health risk would have an estimated value or rating determined by the product:

Risk = Release * Exposure * Exposure-Response * Consequence

where the four variable on the right-hand side are the factors just described. The conceptual units of risk are change in adverse human health consequences per year (or per capita-year, for individual risks) in the exposed population from the proposed change in animal drug use. Of course, this product is most appropriate for a single ***combination*** of the release, exposure, exposure-response, and consequence factors, and hence for a specific animal drug, bacterium, strain (susceptible or resistant), food commodity, exposed susceptible sub-population, and adverse effect category. To estimate or rate total risks, it is necessary to sum the risks over all combinations in the intended scope of the risk

assessment. Thus, multiplicative aggregation of component ratings or estimates is natural for each combination, while additive aggregation is natural across combinations.

Technical Note: Combinations may be thought of as cells in a large contingency table (or as leaf nodes in a classification tree) of factor combinations determining expected illnesses per capita-year for exposed individuals. Given the number of individuals in each cell (its "*size*") and the estimated expected illnesses per capita-year for individuals in that cell (its "*risk*" rate), the expected total illnesses per year in the population is the sum over all cells of the *size*risk* product. The entire probability distribution of total illnesses will be approximately Poisson, and hence determined by the expected number of illnesses. The sum-of-products framework is useful for uncertainty analysis, as products of uncertain factors tend to be approximately log-normal, sums of uncertain products are approximately normally distributed (Cullen and Small), and sums of products may be insensitive to specific numbers (Henrion et al., 1996).

Formal Analysis Results for Possibilities and Impossibilities of Rating Systems

While several qualitative and semi-quantitative risk rating techniques have been developed, as discussed above, there has been little formal analysis of how well they accomplish their intended goals, nor even of what the specific, measurable goals and performance criteria for risk rating systems should be.

Mathematical analysis can help identify the limitations of what any risk rating or risk ranking system can achieve. For example, suppose that a rating system is to be used to compare two different situations, AAUs, or decision options, A and B, to determine which should be ranked higher, e.g., in competing for scarce risk-management resources or in a priority order for regulatory concern and/or intervention. If the overall rating of risk is to be based on component ratings developed for several risk components or factors, as in all of the above examples, then the following abstract analysis may be useful.

For simplicity, suppose there are three component ratings, although the following analysis holds for any greater number of components, not just three. For example, the components might represent Hazard, Exposure, and Impact ratings, as in the Australian system; Probability of exposure, Frequency (or conditional probability) of response given exposure, and Severity of response ratings, as in the Canadian system; or Access, Expression, and Damage scores in the Brenner system. How should the overall risk rating of alternatives A and B depend on the component ratings? Some apparently reasonable properties might include the following.

Possible Desiderata for Aggregating Component Scores into Final Risk Scores

1. Which of alternatives A and B is rated higher in the overall risk rating should depend *only* on their component ratings. Thus, the components used to rate risk should be sufficient to do the job: together, they should determine whether A is assigned a higher, equal, or lower rating than B.
2. Which of A and B is rated higher on overall risk should be able to depend on *each* of their component ratings. Specifically, if A and B are identical in all respects except that A rates higher or worse than B on one factor (e.g.,

exposure), then B should not be rated higher than A in the overall risk rating. This property should hold for all the risk components: none of them is irrelevant.

3. If A rates higher (or worse) than B on *every* component rating, then B should be rated no higher (or worse) than A in the overall risk rating. For example if A involves greater exposure, more illnesses, and more severe consequences than B, then A should receive a risk rating at least as high as B's.
4. Risk ratings of A and B should be based *only on their own data,* i.e., whether A is rated higher or worse than B should not depend on what other alternatives (other than A and B) are also being rated, if any.
5. If one or more component ratings are zero (e.g., for exposure potential or for human health impact potential of exposure), then the overall risk rating should be zero (or "Negligible" in systems with that category).
6. If the rating for a component is uncertain (e.g., if it has a 0.2 probability of being "L", 0.5 probability of being "M", and 0.3 probability of being "H"), then the single "equivalent" rating assigned to it (i.e., H, M, or L after considering its uncertainty) should not depend on the ratings assigned to the other components.

Although such logical relations among the component ratings and the overall risk rating may be desirable, they impose strong constraints on possible rating systems that satisfy them. For example, if quantitative ratings are used, then conditions such as 5 and 6 imply that the aggregation formula used to combine component ratings into an overall risk rating must be *multiplicative,* i.e., the overall risk rating is proportional to a product of its component ratings (http://faculty.washington.edu/jmiyamot/jmfiles/cmb_2col.pdf.) Such multiplicative aggregation of quantitative ratings satisfies properties 1-4. On the other hand, if only qualitative rankings are used for the components, then it turns out that there is *no* qualitative ranking system that can assign coherent overall risk rankings (meaning complete, transitive rank-orderings with ties allowed) based on arbitrary component rank-orderings in such a way that principles 1-4 are satisfied. Similar limitations may hold for aggregating fuzzy ratings of linguistic labels or scales (e.g., H, M, L, and N), depending on how they are formalized (http://www.ie.boun.edu.tr/~taner/publications/papers/ejor.pdf). In other words, qualitative component ratings may not contain enough information to be coherently aggregated into an overall qualitative risk rating that is related to them in desirable ways.

Another possible concern is that a risk rating system with only a few possible outcome categories may not produce enough information to make a good decision if it is *not complex enough* to support effective decision-making. For example, the Australia 3 x 4 matrix assigning a label of H, M, L, or N to each of three components (Hazard, Exposure, and Impact) can provide only a small amount of information (technically, at most six bits of information, equal to the information content of six tosses of a fair coin) to guide decision-makers. Of the much larger quantities of potentially useful and relevant information collected and entered into such a rating scheme (several hundred bits at a conservative estimate), almost all is lost in aggregation during the rating process. The

small fraction that remains (6 bits in this case, or even less if the probabilities of the 12 cells are not all equal) may be insufficient for effective decision-making, which typically requires at least enough information to discriminate among alternatives that have very differently preferred outcomes. The minimum amount of complexity and information required for a classification system (including a risk rating system) to make few errors can be rigorously analyzed via techniques from information theory and computational learning theory (see e.g., Goldman, Chapter 7 and Burges, 1998). A key insight from such formal analysis is that a classification system that lacks enough complexity to discriminate well among essentially different situations may lead to poor decisions, i.e., ratings with high error rates and high expected losses from decision errors.

Rather than further considering properties and limitations of risk rating systems in the abstract, we next focus on constructive approaches for achieving the goals for risk analysis and risk rating systems, building on the multiplicative framework advocated above.

RRRT: A RAPID RISK RATING TECHNIQUE FOR HUMAN HEALTH RISK AND BENEFIT ASSESSMENT

Appendix B develops an example of an algebraic model for assessing the human health risks and benefits of a ban on antibiotic use in animals that affects a foodborne pathogen. Its two key formulas are as follows:

EQUATION A: Expected direct human health benefit from ban = $[p(1 - s)(P^-)MN]*[fr(Q_r - Q_s)]$ illness-days prevented per year by reduced antibiotic resistance in foodborne pathogen = (expected resistant cases prevented per year)*(expected health consequence per case prevented).

EQUATION B: Expected human health harm from ban = $[\Delta F(P^+ - P^-)]MN*[Q_r + s(Q_s - Q_r)]$ illness-days per year in the human population, from reduced antibiotic prevention and control of animal bacterial diseases = (expected cases caused)*(expected consequence per case).

Table 5 summarizes the interpretations and estimated values of the model parameters in these formulas, as calculated in Appendix B. (Intervals indicate ranges of values, roughly interpretable as subjective 95% probability intervals around the point intervals, for use in uncertainty analysis.) The bottom two rows evaluate the above formulas for these parameter values. The formulas can be applied to each animal antibiotic, meat commodity, pathogen, and human population affected by a proposed ban or other change in animal drug use, and the results summed to obtain total population risks and benefits from the intervention is expected to cause. A ban is expected to protect human health if and only if causes a risk reduction that exceeds the risk increase that it causes.

The above two formulas have the following simple intuitive interpretations. The quantity (MN) gives the expected number of chicken servings ingested per year, while the product $(P^-)MN$ is the current expected number of resulting illnesses under the *status quo* (no ban), where (P^-) is the current average risk of illness (i.e., expected number of illnesses caused) per serving. Each illness currently has probability $(1 - s)$ of being resistant to the human antibiotic being considered (e.g., erythromycin) under current animal antibiotic use conditions. However, a fraction p of these currently resistant chicken-caused illnesses, called the *preventable resistance fraction*, would be prevented (i.e., replaced with susceptible rather than resistant pathogens) if the current animal antibiotic use ceased. Thus, the total number of resistant illnesses per year that would be prevented by a ban is: $(P^-)MN(1 - s)p$. Suppose that a fraction f of these would have experienced reduced treatment effectiveness or treatment failure due to resistance if treated with an antibiotic that they are resistant to, and that r is the probability of being treated with a resisted antibiotic. If the mean health impact is $(Q_r - Q_s)$ additional illness-days (or quality-adjusted life-years (QALYs) lost, etc.) for each such case, then each resistant case prevented confers an average health benefit of $f*r*(Q_r - Q_s)$, yielding the final formula Equation A for the total expected direct benefit from a ban.

TABLE 5: PARAMETERS FOR RAPID RISK RATING TECHNIQUE (RRRT) HEALTH RISK-BENEFIT ASSESSMENT MODEL

Symbol	Meaning	Baseline value and source
N	Number of people in population	292E6 (U S Census)
M	Average number of servings of food commodity ingested per capita-year	38 FDA-CVM, 2001, Cox and Popken, 2002 for fresh chicken
P^-	Average probability of severe (treatable) illness per serving from animals without disease. Includes indirect effects of cross-contamination of other foods. This probability is an average for the whole population; individual risks may vary.	1.6583E-8 = (total severe *C. jejuni* illnesses per year)*(fraction caused by chicken)/(total chicken servings ingested per year). See Appendix B.
$P^+ - P^-$	Excess probability of illness per serving from animals with disease (AS+ flocks). (Includes cross-contamination effects)	1.19E-4 (for linear no-threshold dose-response model with microbial load ratio = 10, from Russell, 2003)
ΔF	Fractional change in prevalence of animals with untreated diseases if ban is implemented (and farmers substitute other treatments that are 50% as effective as the banned antibiotic)	0.5% (see Appendix B)
$1 - s$	Fraction of the cases caused by bacteria in animal meat that are resistant to human antibiotic. (s = current *susceptible* fraction)	Erythromycin: 0.01
p	Preventable resistance fraction = fraction of currently resistant illnesses caused by eating the food commodity that a ban would remove (i.e., make susceptible)	Erythromycin: ≤ 1
Q_s	Average human health harm (e.g., days of illness or QALYs lost) per susceptible case. Interpreted as "severity" of a case.	6 days (Marano et al., 2000)
$Q_r - Q_s$	Average excess human health harm (e.g., extra days of illness) per resistant case failing to respond normally to antibiotic, for patients; or per untreated case for non-patients	2 days (Estimated bound for current clinical practice (Ang and Nacham, 2003);
f	Probability that resistant case fails to respond normally (i.e., the same as a susceptible case) to assigned antibiotic therapy due to resistance	Erythromycin: ≤ 1 [0, 1]
r	Probability that a resistant case is assigned resisted antibiotic	0.5, [0.25, 1]
Risk prevented	$\leq$ **1.84** illness-days/yr. for macrolides	$p(1 - s)fr(Q_r - Q_s)(P^-)MN$
Risk created	**40,458** = 0.005*1.19E-4 *[8 – 0.9362*2] *38*292E6 excess illness-days per year = 6602 additional cases*6.13 days/case.	$[\Delta F(P^+ - P^-)]*[Q_r + s(Q_s - Q_r)]MN$

Equation B is equally interpretable. If a ban would cause an increase ΔF in the fraction of chicken servings from airsacculitis-positive (AS+) flocks instead of airsacculitis-negative (AS-) flocks, and if each such serving has an incremental probability $(P^+ - P^-)$ of causing illness, then the expected change in illnesses will be $[\Delta F(P^+ - P^-)]MN$. Suppose that a faction s of these illnesses are susceptible and that the average health impact per illness caused is therefore $[sQ_s + (1 - s)Q_r]$, which may be rearranged as $[Q_r + s(Q_s - Q_r)]$. Then the expected human health impact caused by the change ΔF in animal illness prevalence is: $[\Delta F(P^+ - P^-)]MN*[Q_r + s(Q_s - Q_r)]$ illness-days (Equation B). Equation A expresses the expected human health benefits from treatment failures prevented by

reducing resistance, while Equation B expresses the expected human health harm caused by reduced prevention of animal illnesses and associated microbial loads in processed foods. Thus, Equation A is applied to the subset of campylobacteriosis patients who might benefit from an antibiotic, i.e., those with relatively severe cases (Ang and Nachman, 2003; http://www.cdc.gov/ncidod/dbmd/diseaseinfo/campylobacter_g.htm); while Equation B applies to the wider population of people who eat chicken.

Some of the parameters in Table 5 are not directly measured, but must be calculated from other data elements. Appendix B gives the details of the data and parameter estimate calculations and discusses the judgments and assumptions used to select approximate uncertainty factors or uncertainty intervals where data are inconclusive. For convenience, Table 4 summarizes the main data elements, calculations of parameter point estimates, and model results. Table 5 presents a perhaps more logical, simple and uniform format for summarizing the key quantities and calculations for decision-makers, as recommended by Bailar and Travers, 2002, whereas Table 4 gives greater visibility into the data elements and sources used to calculate the quantities in Table 5.

Together, Tables 4 and 5 illustrate some of the possibilities for rapid, well-documented risk assessments made possible by implementing the multiplicative risk assessment approach suggested by Bailar and Travers, 2002.

EXAMPLE RRRT CALCULATIONS FOR COMMENSALS: VIRGINIAMYCIN RISK

To further illustrate the application of the RRRT risk assessment framework, Table 6 applies it to estimate human health risks from streptogramin-resistant vanA vancomycin-resistant *E. faecium* ($VREF_A$) due to the continued use of the streptogramin combination quinupristin-dalfopristin (QD, marketed as virginiamycin (VM) for animals and as Synercid™ for human patients). This application is considered further at the end of the Hazard Identification section. Details of the calculations and data are in Cox and Popken, 2004. For purposes of illustrating the RRRT approach, the main point is that the calculation of *status quo* human health risks for commensals and for the at-risk population (ICU patients with compromised immune systems) can again be structured as a product of a base rate of illnesses with a series of fractions, all of which can either be estimated from data or set equal to one as a default upper bound. Table 7, taken from Cox and Popken, 2004 compares model parameter values in the US and Australia. It also shows the probability distributions used to estimate the uncertainty distribution of risk via Monte Carlo simulation. The bottom rows of Table 7 indicate how to model a gradual decline over time in QD prescription rates as Synercid™ is increasingly replaced with linezolid (Zyvox™), and a gradual decline in QD resistance in chickens (and, presumably, a proportional decline in QD resistance among chicken-borne SREF cases in humans) following a withdrawal of virginiamycin. In other words, the multiplicative framework need not be confined to comparing pre-intervention and post-intervention steady-state equilibria, but can be extended to consider the transient adjustments in factors caused by interventions over a time horizon of years or decades.

TABLE 6: RRRT Estimation of Health Risks from Virginiamycin Use in Chickens

PREVENTABLE EXPOSURE TO RESISTANT BACTERIA VIA FOOD		
Factor	**US Values**	**Data Sources**
Average number of VRE cases/year in ICU population	37,483	NNIS, 2001 Lawton et al, 2000 AHA, 2001
Fraction of VRE cases that are VREF	0.71, **0.78**, 0.95, Median: 0.78, UF < 1.25	SNJ, 2000 Clark et al., 1993 Rice, 2001
Fraction of VREF cases with VanA resistance ($VREF_A$)	0.73 0.83 **0.79** Median: 0.79, UF < 1.1	Eliopoulos et al., 1998 Jones et al, 1995 Clark et al., 1993
Fraction of $VREF_A$ cases from food	≤ 0.17 = Proportion of VREFs that are not of known nosocomial origin	Bischoff et al., 1999; Austin et al., 1999; Thal et al., 1998
Fraction of $VREF_A$ cases from food that might come from chickens	0 to 0.12 based on genogroup similarities	Willems et al., 2000 Willems et al., 2001
Fraction of foodborne $VREF_A$ cases that have QD-resistance (= $SREF_A$ cases)	0 to 0.011	Eliopoulos et al., 1998, Jones et al., 1999
Fraction of foodborne $SREF_A$ cases with QD-resistance caused by QD use in chickens	1 (?)	Upper bound
Preventable resistance fraction = fraction of foodborne $SREF_A$ cases that could be prevented if QD use in animals ceased	≤ 1. Cox and Popken, 2004 estimate 0.68 within 5 years based on Danish experience in chickens	Upper bound
CONSEQUENCES OF QD-RESISTANCE		
Fraction of $SREF_A$ cases not treated successfully with linezolid or other non-QD antibiotics	0.074 = fraction of cases for which linezolid therapy is not successful	Linden et al., 2002
Fraction of $SREF_A$ cases not treated successfully with linezolid or other non-QD antibiotics that are then prescribed QD	1 (?)	Upper bound
Fraction of $SREF_A$ cases prescribed QD that fail to respond normally to QD treatment	1 (?)	Upper bound
Fraction of $SREF_A$ cases prescribed QD that fail to respond normally to QD treatment *because of* the QD-resistance	0.7	Linden et al., 2002 Moellering et al., 1999
Increased mortality probability due to QD resistance	0.15	Linden et al., 1997
Preventable excess mortalities per year = 0.04 = 37483*0.78*0.79*0.12*0.011*0.17*0.074*0.7*0.15	≤ 0.04	Product of above upper bounds

TABLE 7: Comparison of US and Australia Parameter Values for $SREF_A$ Risks

Process	**Formula** (Australia; US)	**Mean** (Australia)	**Mean** (US)
VRE cases/quarter	Markov Simulation Model; 13.84 x N(677, 22.34)	3.98	9370.65
Time Independent Reductions			
Van A *E. faecium* prop.	Beta(18,65); Uniform(.43,.79)	0.22	0.61
Exogenous case prop.	Uniform(0.089, 0.25)	0.17	0.17
Chicken attribution prop.	Beta(11, 78)	0 to 0.12	0 to 0.12
QD resistance prop.	Beta(1, 109); p = 6/553	0 to 0.009	0 to 0.011
QD treatment effectiveness proportion	N(.705, .0362) + Beta(1,109); N(.705, .0362) + Bin(6/553)	0.714	0.716
Summary of reductions	Product of above	**.000029;**	**.0001**
Time Dependent Reductions	(t represents quarters)		
QD prescription rate	Decrease 15% semiannually	0.922^t;	$0.922^{(t+6)}$
VM resistance reduction in chickens after ban	Decreases to 0.32 after 5 years	$e^{(-.0570\ t)}$	$e^{(-.0570\ t)}$

Source: Cox and Popken, 2004.

The risk assessment summarized in Table 6 assumes that hospitals with outbreaks of nosocomial $VREF_A$ or $SREF_A$ transmission are approximately equally likely to have them whether or not VM is used in chicken (based on empirical evidence that chicken-borne infections are at most a relatively minor contributor, reviewed in Cox and Popken, 2004), so that only non-nosocomial cases are considered as being potentially preventable by reduced VM use in food animals. In contrast to the analysis in Table 7, Table 6 also assumes that QD is used to treat a patient with a vanA infection only if other antibiotic treatment options, specifically linezolid, prove ineffective. (Since these patients are typically seriously ill and carefully screened and monitored, it is much less likely that QD would be prescribed to patients with QD-resistant infections than that macrolides would be prescribed as empiric treatments for macrolide-resistant *C. jejuni* cases in Table 4.) With these assumptions and the point estimates of model parameters shown in Table 6, an upper-bound point estimate for preventable excess mortalities per year from QD use in animals is about 0.04 cases per year, or about 1 excess mortality in 25 years. The true number of excess deaths could be far smaller (possibly zero), as a number of conservative (risk estimate-maximizing) assumptions are included in calculating the 0.04 value in Table 6.) These statistical excess deaths occur among seriously ill patients, over $1/3^{rd}$ of whom are expected to die even if there is no resistance.

By contrast, the human health benefits from continued use of virginiamycin in chickens may be comparable in magnitude to those estimated for continued use of macrolides in Table 4. VM is effective in reducing the incidence of necrotic enteritis (NE)

in chickens. Withdrawing VM may lead to increased NE rates (consistent with experience in Europe, where NE rates in chickens experienced a transient surge and new, higher endemic levels in several countries following the withdrawal of VM and other antibiotics used as animal prophylactics and growth promoters.) Chicken carcasses from NE-positive (NE+) flocks are likely to have the same kinds of excess microbial loads observed for carcasses from AS+ flocks (Russell, 2003). They may lead to similar excess human illnesses, i.e., about 40,458 excess illness days and 39.3 serious *C. jejuni* illnesses per half percent increase in NE+ flocks if virginiamycin were withdrawn, based on the numbers at the bottom of Table 4 for airsacculitis. In short, a withdrawal of virginiamycin from use in chickens may create substantial adverse human health impacts – possibly greater than those estimated here if human health effects of other chicken-borne pathogens (e.g., Salomonella) that might be affected by increased NE are considered. Without pursuing the benefit calculation further, it seems clear that uncertainty about the extent to which NE would increase after a risk management intervention, and the extent to which such an increase causes increases in chicken carcass contamination and human illnesses, are worth additional empirical investigation. Data from Europe may be valuable for this purpose (e.g., Eurosurveillance, 2002).

METHODS AND DATA FOR RISK ASSESSMENT

The previous RRRT example calculations have illustrated many of the types of top-down calculation methods and data sources used in antimicrobial risk assessment. More generally, quantitative risk assessments are often simplified by applying results from probability and statistics. Among the most useful are expressions for the probabilities of conjunctions of events as products of marginal and conditional probabilities (extensively used in the RRRT approach) and "limit laws" that allow the probability distributions of population risks to be closely approximated based on partial knowledge of the probability distributions of the factors that contribute to them. For example:

- *Rare events* typically obey a Poisson approximation law (e.g., Barbour, 2000)
- *Sums and averages* (e.g., total population risks or average individual risks) of independent or almost independent variables (e.g., individual risks) typically approach normal distributions in large populations.
- *Products and networks* of calculations often give results with approximate log-normal distributions (Druzdzel, 1994).
- *Extreme values* (e.g., maximum or minimum values, record values) from a large population usually follow one of three kinds of extreme value distributions. Runs of large or small values also follow special distributions.
- *Deviations* around expected values often follow an approximately exponential distribution.

Such results can allow population risks to be approximated with useful accuracy for large populations and complex models even when there is considerable uncertainty about the values (or probability distributions) of individual factors in the models. For further discussion, see Appendix A.

HAZARD IDENTIFICATION

Risk assessment begins with *hazard identification*, which defines the *scope* of the assessment – what specifically will be assessed – and presents evidence that a particular activity or source of risk (the hazard) causes harm in exposed individuals or populations.

DEFINITION OF HAZARD IDENTIFICATION

Hazard identification for food safety has been defined as "The identification of biological, chemical and physical *agents capable of causing adverse health effects* and which may be present in a particular food or group of foods" (Codex Alimentarius Commission, http://www.fao.org/DOCREP/005/Y2200E/y2200e07.htm, emphasis added). These agents, i.e., sources of risk, are called *hazards.*

A ***hazard*** is thus a *potential cause of an adverse human health effect.* Examples may include food-borne bacteria and resistance determinants transferred from food-borne bacteria to other infectious bacteria that cause them not to respond to treatment, creating increased days of illness or other clinical harm. Potential ***adverse human health effects*** (or ***consequences***) of exposures to hazards could include increased frequency, duration, or severity of food-borne illnesses, or treatment failures that result in clinical harm (e.g., increased duration or severity of illnesses).

PURPOSES OF HAZARD IDENTIFICATION

Hazard identification has the following main purposes:

1. *Rapidly screen potential hazards* by identifying whether available data support the hypothesis that specific adverse health effects might be caused by specific exposures or activities. Hazard identification uses methods of causal analysis (e.g., Shipley, 2000) to determine whether hypothesized causal relations relating acts to exposures to adverse health responses and consequences are consistent with available data.

2. *Identify qualitative or quantitative causal relations between exposures to specific food-borne hazards and specific adverse human health effects.* To support risk management decision-making, it is often helpful to identify exposures or hazards resulting from controllable decisions or behaviors.

3. *Identify risk factors and exposure conditions that are associated with increased risks to specific exposed populations* (e.g., the old, the young, the immuno-compromised, etc.)

4. *Present and objectively evaluate evidence for and against the hypothesis that exposures to specific food-related hazards (resulting from controllable decisions, e.g., on use of feed additives) cause specific adverse human health effects.* This is

somewhat analogous to the US EPA's statement that, for environmental hazards, "The objective of hazard identification is to determine whether the available scientific data describe a *causal relationship* between an environmental agent and *demonstrated injury* to human health or the environment" (http://www.bethel.edu/~kisrob/hon301k/readings/risk/RiskEPA/riskepa1.html, emphasis added).

If the hazard is non-zero, what agent should be considered "the cause" of adverse consequences – the bacteria involved, the resistance determinants that they carry, failure to properly prepare, cook, or handle food to eliminate contamination, or the animal antibiotic uses that selects for those determinants? The answer depends on what risk management decision the hazard identification is intended to support. In general, "the cause" of an adverse health effect is not uniquely defined, but the predicted effect of a specific intervention, holding other conditions fixed, can be quantified.

DATA USED IN HAZARD IDENTIFICATION

Objective statistical tests for potential causality between measured exposure and response variables can be applied if historical data are available on the timing and/or extent of the variables in a population. For example, suppose that historical data are available for the following two variables:

- X = animal antibiotic use in a country or location (perhaps coded so that X(t) = 1 if the antibiotic was in use at time t, else 0, or coded so that X(t) indicates the levels of use at time t); and
- Y = resistance rates to analogous human antibiotics in human patients (e.g., from surveillance data or epidemiological studies).

Then X is a *potential cause* of Y if and only if the history of X up to any date t provides information about the future of Y (after date t) that cannot be fully removed by conditioning on any other subset of variables known at date t (including the history of Y itself up through date t). Whether one variable provides information about another (indicated statistically by a non-zero cross-entropy between them, i.e., a reduction in the entropy of one after conditioning on the other) and whether this information can be removed ("explained away") by conditioning on other variables can be determined by computational statistical algorithms (Frey et al., 2003; Aliferis et al., 2003).

Categories of objective evidence that are often considered in antimicrobial risk assessments include:

- *Spatial associations* between animal antibiotic use and resistance levels in human patients. Associations between animal antibiotic use and human illness rates may also be relevant if the animal antibiotic use affects microbial loads of pathogens reaching human via meats.
- *Temporal associations* between the date(s) of *introduction* of an animal antibiotic and subsequent changes in animal and human resistance rates (after controlling for

contemporaneous changes in other factors and potential confounders, e.g., changes in foreign travel).

- *Temporal associations* between the date(s) of *cessation* of an animal antibiotic (e.g., following the European bans on growth promoters) and subsequent changes in animal and human resistance rates (after controlling for contemporaneous changes in other factors, e.g., consumer awareness and education programs, HACCP interventions).
- *Genetic associations* between bacteria found in human patients and in food animals that may indicate whether they are similar enough so that one might come from the other (or whether both might have a common environmental source). Usually, epidemiological data are invoked to help interpret and complement genetic similarity data, since genetic similarities alone cannot establish a direction of causation.
- *Epidemiological associations* between exposures to food animal products and incidence rates of foodborne illnesses and/or prevalence rates of resistance in patients, after controlling for potential confounders, information biases, and modeling biases.

Well-developed statistical methods and algorithms are available to identify significant statistical associations from such relevant data and to screen them for potential causality based on the above information criteria.

Technical Note: Statistical tests for potential causality. Statistical methods are available to identify associations that cannot be "explained away" by conditioning, even in very large data sets (Aliferis, 2003). These algorithms require tests for conditional independence as sub-routines. Classification tree software can be used to perform conditional independence tests for one dependent variable at a time by testing whether conditional mutual information is significantly different from zero (Frey et al., 2003). Alternatively, statistical tests of the residuals in flexible nonparametric ("form-free") regression models (Linton and Gazalo, 1999; Shipley, 2000) can be used to test conditional independence for one dependent variable at a time. More computationally-intensive commercial software (e.g., BayesiaLab™) will automatically compute conditional independence relations for entire sets of variables (Tsamardinos et al., 2003). These algorithms generalize the requirement that, to be considered causal, an exposure-response association must *not* be fully explained by confounding (Sonis, 1998; Greenland and Morgenstern, 2001; Greenland, 2003) – or, for that matter, by sample selection biases (Mark, 1997), information biases (Grimes and Shulz, 2002), or modeling and analysis biases (Cox, 2001). Formal tests for statistically significant associations between the timing of one event (e.g., introduction or cessation of animal antibiotic use) and subsequent changes in a series of measurements (e.g., human resistance rates in a surveillance program) can be based on *intervention analysis* and *change point analyses* (Green, 1995) for time series. Potential causality between two time series of measurements (e.g., usage levels of an animal drug and illness rates or resistance rates in human patients) can be based on extensions of *Granger-Sims tests* (Swanson et al., 2001) that include conditional independence and causal graph tests. These methods represent the current state-of-the-art in causality testing. They are entering common biostatistical and risk analysis practice only slowly, but have been developed for many decades in other disciplines (Shipley, 2000).

DESIRED OUTPUTS OF HAZARD IDENTIFICATION

A hazard identification for microbial risk assessment should identify the microorganism that causes specific diseases or adverse health effects (e.g., using Koch's

postulates), elucidate the infection and disease process (including the conditions under which infection and illness occur); identify possible transmission routes (e.g., food, water, vectors); and identify covariates (e.g., host immune status, other risk factors) that can interact with or affect the relation between exposure and risk (Haas CN, Rose JB, Gerba CP. *Quantitative Microbial Risk Assessment.* Wiley, 1999). A hazard identification for antimicrobial risk assessment should also identify the causal relation between use of antimicrobial drugs in animal feed additives and the levels of resistant pathogens in human patients, as well as the causal relation between these levels and the frequency and magnitude of increased mortality risks, morbidity risks, and treatment failure rates.

If objective statistical tests for hazard identification do not identify a causal relation between decisions, exposures, and human health risks, then this result should be stated, along with discussion of the statistical power of the tests used for the data examined. In this case, risk assessment can still be carried out, but it becomes contingent on the assumption that a risk exists. Such a contingent risk analysis can be useful if it shows that risks are small, by providing a plausible upper bound on the true (conjectured but perhaps non-existent) risk. But it may not be useful for accomplishing other risk analysis goals, such as guiding rational choice among expensive risk management alternatives.

EXAMPLES OF HAZARD IDENTIFICATION

Example 1: Sources of antibiotic-resistant E. faecium

It is an *a priori* plausible hypothesis that antibiotic-resistant *E. faecium* isolated from human patients might originate in antibiotic-treated food animals that carry *E. faecium* (Wegener et al., 1999). The hazard identification step of the risk assessment process rigorously tests and evaluates such hypotheses using data. It may also use epidemiological, time series, genotype, phenotypic biomarker, and other mechanistic data to investigate sources of exposure to microbial hazards (e.g., antibiotic-resistant *E. faecium*) even without any *a priori* hypotheses.

How can microbiological hazard identification methods be applied to identify sources of antibiotic-*resistant E. faecium*? One approach is to compare resistance phenotypes in isolates from different sources. This is illustrated by a study of Iversen et al. (2004):

> "An ampicillin- and ciprofloxacin-resistant *Enterococcus faecium* (ARE) strain, named FMSE1, with a characteristic biochemical phenotype, was in a recent study found to dominate among faecal ARE isolates from patients in several Swedish hospitals. In the present study, the prevalence of this strain among 9676 enterococcal isolates from healthy children, hospital sewage, urban sewage, surface water, slaughtered animals (broilers, pigs and cattle) and pig faeces and manure was investigated. Enterococcal isolates having the same biochemical phenotype as the FMSE1 were most common in samples of hospital sewage (50%), surface water (35%), treated sewage (28%) and untreated sewage (17%), but rare in samples from healthy children (0.8%) and animals (2%). PFGE typing of FMSE1-like isolates from hospital sewage indicated that they were closely related to the nosocomial FMSE1 strain. Thus, this study indicated a possible transmission route for nosocomial E. faecium from patients in hospitals to hospital sewage and urban sewage, and further via treatment plants to surface water and possibly back to humans. This proposed route of circulation of drug-resistant enterococci might be further amplified by antibiotic usage in human medicine. In contrast, such transmission from food animals seems to play a negligible role in Sweden."

Such studies can help to identify hazards that were not necessarily expected *a priori*. Conversely, as in this case, they can help show the extent to which potential hazards that might seem plausible in the absence of data, such as the foodborne transmission pathway, make a significant contribution to human illness in reality. The conclusion from this study that "transmission from food animals seems to play a negligible role in Sweden" might not have been anticipated in the absence of data-driven hazard identification, as many scientists have taken as axiomatic the assumption that foodborne transmission plays a major or predominant role in human resistant illnesses (e.g., FAAIR, 2002; Wegener, 2003).

Example 2: Potential human health hazard from an existing animal antibiotic use – Tylosin

The macrolide Tylosin is used in soluble and premix formulations in chickens to prevent and control several bacterial diseases and to promote health and growth. Tylosin use in chickens can potentially affect human health by changing microbial loads on chicken products and/or by selecting for macrolide-resistant pathogens and commensals reaching people on food commodities. Food-borne bacterial pathogens found in chicken that are of specific concern as potential hazards to human health are *C. jejuni* and *C. coli,* both of which are found in live broilers (Wedderkopp et al., 2003), chicken carcasses, and retail chicken products (Ge et al., 2003; Musgrove et al., 2003). Diagnosed cases of severe campylobacteriosis in humans may be treated with erythromycin or other macrolides. In addition, if tylosin use in chickens selects for macrolide-resistant *E. faecium* that are also streptogramin A-resistant, then these multi-drug resistant *E. faecium* might increase the risk from serious infections in patients with compromised immune systems. Finally, the potential for induction, selection, and transfer of resistance determinants to other bacteria that infect humans must be assessed.

Although macrolides are important antibiotics in human medicine, this is relevant for risk assessment of tylosin use in chickens only to the extent that such use reduces the effectiveness of macrolides used in human medicine. A human health risk exists only to the extent that there is potential to cause harm to human health, and not simply as a result of macrolides being important in human medicine. The *human health risk* from tylosin use in chickens may be defined as the *expected number of additional illness cases per year* (for population risks) and per capita-year (for individual risks) *in each illness severity class* (mild, moderate, severe, fatal) caused by tylosin use in chickens, along with the resulting annual frequency distribution of consequence severities, i.e., excess annual illness-days and quality-adjusted life years lost in each severity class. Unless this number is positive for one or more severity classes, the hazard is zero.

A current clinical perspective on the treatment of *C. jejuni* infections is as follows:

> "Most *C. jejuni* infections are mild and self-limited; therefore, they do not usually require antibiotic therapy. Correction of electrolyte abnormalities and rehydration are usually sufficient. Treatment often is reserved for compromised hosts or persons with fever, increasing bloody diarrhea, or symptoms that last longer than 1 week. *C jejuni* is usually sensitive to erythromycin, gentamicin, tetracycline, ciprofloxacin, and clindamycin. Reports of erythromycin- and ciprofloxacin-resistant strains are increasing. In adults, placebo-controlled studies of erythromycin demonstrate no improvement in the clinical symptoms if given late in the course of illness but have resulted in decreased fecal shedding. If an appropriate antibiotic therapy was initiated within the first 4 days of illness, there was a reduction in the excretion of the organism; however, results regarding the resolution of symptoms were controversial. In contrast, early erythromycin treatment for children with bloody diarrhea shortened both the duration of diarrhea and excretion of microbes in the stool. Recommended duration for antibiotic treatment given for gastroenteritis is 5-7 days. Antimicrobial therapy for all bacteremic and immunocompromised patients with *C jejuni*

should be selected based on a laboratory susceptibility test. Begin therapy with gentamicin, imipenem, third-generation cephalosporins, or chloramphenicol until susceptibility test results are available." (Ang and Nachman, 2003)

This perspective is consistent with clinical experience gathered over the past twenty years. For example, on the issue of clinical effectiveness, despite some initial promising reports on the efficacy of erythromycin in shortening the duration of *C. jejuni* campylobacteriosis (Nolan et al., 1983), others soon found that "Although erythromycin significantly shortened the duration of *C jejuni* excretion, it appeared to exert no effect on the clinical course of the illness" (Robins-Browne et al. (1983); see also Anders et al., (1982)). When investigators focused specifically on early treatment, they still found that that "Erythromycin rapidly eliminated *C. jejuni* from [human] stools.... Despite its bacteriologic effectiveness, erythromycin did not reduce the duration or severity of diarrhea, abdominal pain, or other symptoms" (Williams et al. 1989).

Since macrolides (e.g., erythromycin or azithromycin) are not drugs of first (or second or third) choice for high-risk patients, have limited or no clear clinical benefits in adults, and have many alternative products available, it might at first seem that their overall importance in human medicine for treating *C. jejuni* infections is very limited. However, if macrolides do after all have some (perhaps currently unrecognized) clinical benefits in treating *C. jejuni* cases, then animal uses that select for resistant strains may cause some or all of those benefits to be lost. Concern that this could be the case has driven risk management recommendations that animal antibiotic use be terminated. However, such recommendations have not been based on formal risk assessment or hazard identification showing that macrolide use in chickens explains part of the observed resistance patterns in human patients.

As discussed further in Appendix B, recent historical data that could potentially have provided evidence of a causal relation between macrolide use in food animals and macrolide resistance rates in human patients instead produced some surprising findings. Specifically, the European ban on macrolides and other antibiotics used as prophylactics and growth promoters in animals (including Tylosin) was followed by unexpected *increases* in macrolide resistance rates in humans as well as in rates of foodborne illnesses in Denmark and some other countries (Hayes and Jensen, 2003), contrasting with previous expectations and opinions (e.g., Wegener et al., 1999; Wegener, 2003). No cause-and-effect relation has yet been established, and the matter is politically and scientifically controversial. In the absence of formal hazard identification and causal analysis of the data, competing explanations have been proposed, e.g., that the bans led to increased therapeutic use (although to decreased total animal use) of macrolides; that food imports may account for increasing shares of observed illnesses; or that the bans increased animal and human illnesses and hence increased human antibiotic use and resulting antibiotic resistance in humans. Further examination of the European post-ban experience may produce additional valuable information on causal relations, and policy-makers may take advantage of this information and of hazard identification methods in determining whether such bans actually cause the human health benefits that they are intended to achieve.

Example 3: Potential human health hazard from an existing animal antibiotic use – Virginiamycin

E. faecium are commensal bacteria, commonly found in the intestines of humans and of food animals such as chickens, pigs, and cattle. Competent immune systems protect most people from *E. faecium* infections. However, patients with compromised immune systems, such as leukemia, chemotherapy, transplant, and AIDS patients, can develop life-threatening *E. faecium* infections unless these bacteria can be controlled successfully with antibiotics. Infections typically

occur in intensive care unit (ICU) patients, usually via nosocomial transmission. Vancomycin is the antibiotic most frequently prescribed to treat *E. faecium* infections, but may be ineffective against *E. faecium* that express resistance genes. Other antibiotics such as linezolid, daptomycin, and quinupristin-dalfopristin, which are usually highly effecive against vancomycin-resistant *E. faecium* (VREF), may then become important treatment options (Critchley et al., 2003). Less effective bacteriostatic agents (e.g., chloramphenicol) are also available, and new antibiotics for treatment of vancomycin-resistant cases (e.g., oritavancin, a glycopeptide, and tigilcycline, a novel analogue of minocycline) are in trial (Linden, 2002).

A current clinical perspective on VREF infections and resistance is as follows:

> "The acquisition of vancomycin resistance by enterococci has had serious implications for the treatment and infection control of these organisms. Vancomycin-resistant enterococci (VRE), particularly *E faecium* strains [VREF], are frequently resistant to all antibiotics that are effective treatment for vancomycin-susceptible enterococci, which leaves clinicians treating VRE infections with either suboptimal bacteriostatic agents (eg, chloramphenicol) or with no therapeutic options. Recently, 2 new types of antibiotics (quinupristin/dalfopristin, linezolid) with activity against many VRE strains have improved this situation, but resistance to both of these agents has already been described. ... Five phenotypes of vancomycin resistance, termed VanA, VanB, VanC, VanD, and VanE, are described. The VanA and VanB phenotypes are clinically significant and mediated by 1-2 acquired, transferable operons consisting of 7 genes in 2 clusters termed VANA and VANB operons. ...In the United States and Europe, the VanA-resistance phenotype is reported as the most common phenotype. VanA enterococcal isolates exhibit high-level resistance to both vancomycin and teicoplanin, while VanB isolates have variable resistance to vancomycin and remain susceptible to teicoplanin. ...Enterococcal infections often occur in debilitated patients and as part of polymicrobial infections. These factors limit the ability of investigators to determine the independent contribution of enterococcal infections to mortality and morbidity. ...Enterococcal infections are more common in elderly patients because of various associated factors that are more common in these patients. ...The streptogramin combination antibiotic, quinupristin/dalfopristin, is available intravenously for the treatment of *E faecium* infections, but it is not effective against *E faecalis* strains. Linezolid, an oxazolidinone antibiotic, is available orally and intravenously, and it is used to treat infections caused by *E faecium* and *E faecalis* strains. ...Once VRE is identified in a medical facility, all clinical enterococcal isolates should be tested for vancomycin resistance." (Donskey and Salata, 2003)

Thus, either the streptogramin combination quinupristin-dalfopristin (QD) or linezolid can potentially be used to treat vancomycin-resistant *E. faecium* (VREF) infections. However, suppose that use of QD (formulated as virginiamycin) in food animals were to increase streptogramin-resistant *E. faecium* (SREF) among VREF in meat products, and hence increase QD-resistant VREF infections in immunocompromised patients, perhaps following inadequate cooking of hospital food. Then more of these patients might have to be treated with alternatives such as linezolid or daptomycin instead of with QD. Since linezolid is usually less harsh and at least as effective as QD, this is not necessarily undesirable. However, for patients who do not respond favorably to linezolid – approximately 7.4% of VRE patients in a study by Linden et al, 1997 – or to other therapies such as daptomycin, QD may become the treatment of last resort. QD resistance may then increase the probability of QD treatment failure. Therefore, if QD use in food animals increases the rate of QD-resistant VREF infections in ICU patients, it might also increase the number of cases per year not treated effectively by any currently available antibiotics, leading to excess mortalities. Quantitative risk assessment is needed to determine how large this number is. As discussed in Example 1 above, identifying an *a priori* plausible causal pathway, such as the foodborne one just described, that might lead from an activity (QD use in food animals) to an adverse human health effect (increased treatment failures among QD-treated patients), does not by itself constitute an adequate hazard identification. Rather, such hypothesized causal pathways must be tested against available data (Shipley, 2000). When data are not adequate to permit such testing

of causal hypotheses, they may be identified as *potential* hazards (not known to be false), and risk assessment can then be carried out *contingent* on the assumption that the conjectured hazards are real, to obtain upper bounds on the true but unknown risks (Cox and Popken, 2004).

EXPOSURE ASSESSMENT

DEFINITION OF EXPOSURE ASSESSMENT

Exposure assessment has been defined as "the qualitative and/or quantitative evaluation of the *degree of intake* likely to occur" (Joint Expert Consultation of the FAO/WHO, http://www.foodriskclearinghouse.umd.edu/pversion/exposure.htm, emphasis added). In terms of the causal chain:

$$\text{act} \rightarrow \Delta\text{exposure} \rightarrow \Delta\text{illnesses} \leftarrow \text{covariates},$$

exposure assessment describes the act $\rightarrow$ Δexposure link. It provides qualitative or quantitative summaries of the population exposures to microbial loads (resistant, susceptible, or both) for different risk management decisions. Exposures for individuals in the population can be expressed in units of frequency and magnitude of CFUs, or colony-forming units, ingested via food, water, from contaminated hands, and via other pathways. For populations, exposure refers to the *frequency distribution of individual exposures* (microbial loads) consumed per unit time.

The US FDA has defined exposure assessment as "A component of a risk assessment that characterizes the source and magnitude of human exposure to the pathogen". The magnitude of human exposure, also called the dose, is defined as "The amount or number of a pathogen that is ingested or interacts with an organism (host)" (http://www.foodsafety.gov/~dms/lmriskgl.html). This is roughly analogous to concepts used in environmental risk assessment. For example, US EPA experts have stated that "Questions raised in the exposure analysis concern the likely sources of the pollutant... its concentration at the source, its pathways (air, water, food) from the source to target populations, and actual levels impacting target organisms."
(http://www.bethel.edu/~kisrob/hon301k/readings/risk/RiskEPA/riskepa1.html).

Consideration of the amount of contamination (i.e., the frequency distribution of microbial loads) ingested by individuals is crucial for quantifying risk. This reflects the fundamental principle that "the dose makes the poison". However, in practice, it is quite common for the microbial loads received to be very uncertain, especially if they depend on unmeasured and/or highly variable processes such as cooking of food, cross-contamination of other foods in the kitchen, or transfer from contaminated surfaces to skin to ingestion. In such cases, the exposure assessment diagram may look like this:

$$\begin{array}{l} \text{act} \rightarrow \text{exposures} \rightarrow \text{illnesses} \leftarrow \text{individual covariates} \\ \qquad\qquad\quad \downarrow \\ \qquad\qquad\quad \text{measured exposure surrogates} \end{array}$$

Available data may consist of surrogate measurements (e.g., microbial concentrations in carcass rinses at retail, or on swabbed surfaces) rather than direct measurements of the ingested microbial loads that cause infection or illness. Exposure assessment with

surrogate exposure measurements consists of estimating how the underlying true exposures will change if different risk management actions are taken, while subsequent exposure-response modeling must focus on how health risks will change when true exposures are changed by decisions. True exposures then play the role of *latent variables* in causal modeling. Appropriate statistical techniques for inferring the relation between decisions and true exposures can still be applied using the above modified diagram with surrogate measurements of exposure for data. Software such as WinBUGS helps to automate the required computations for inference with missing data and surrogate variables.

In summary, exposure may be defined as the number of servings containing potentially infectious doses of the bacterium of concern ingested per year (for population exposure) or per capita-year (for individual exposure). A "potentially infectious dose" is any dose large enough to infect a susceptible consumer. It may be as small as one CFU, if that is biologically realistic. If reliable dose-response information shows that the risk of illness below some number of ingested CFUs is negligible [or, more precisely, is small enough so that it can be ignored without changing the expected number of illnesses per year (within the limits of rounding error)], then a "practical threshold" – meaning one that leads to numerically accurate risk calculations – may be used, even if, in principle, no true biological threshold exists. The number of servings per year ingested with microbial loads above the practical threshold then defines total annual exposure.

PURPOSES OF EXPOSURE ASSESSMENT

Exposure assessment has the following goals:

- Identify exposed sub-populations at risk of infection and illness
- Identify conditions leading to high-risk exposures
- Describe the extent of exposures (= frequency and magnitude of individual exposure in the population in relation to susceptibility/covariates)
- Predict how risk management decision options will affect exposure distribution

DESIRED OUTPUTS OF EXPOSURE ASSESSMENT

A successful exposure assessment describes the frequency distribution of potentially infective microbial loads in exposed populations and sub-populations. It shows how these distributions change for different risk management decisions. The descriptions may be qualitative or quantitative, but they should contain enough variety to indicate any significant differences in microbial load distributions for different decisions.

EXAMPLES OF EXPOSURE ASSESSMENT IN THE RRRT FRAMEWORK

The top portions of Tables 4 and 6 illustrate how exposure calculations are performed within the RRRT framework. In both tables, the product of the first 8 factors gives an estimate of the potentially preventable resistant illnesses per year caused by current animal antibiotic use.

EXPOSURE-RESPONSE AND DOSE-RESPONSE MODELING

DEFINITION OF EXPOSURE-RESPONSE MODELING

Following the National Academy of Sciences framework for risk analysis, the US FDA, CDC and USDA defined dose-response assessment as "The determination of the relationship between the magnitude of exposure and the magnitude and/or frequency of adverse effects" (http://www.foodsafety.gov/~dms/lmriskgl.html). The Codex Alimentarius Commission states that "For biological or physical agents, a dose-response assessment should be performed if the data are obtainable."

PURPOSES OF EXPOSURE-RESPONSE MODELING

Dose-response modeling, or exposure-response modeling, describes the causal relation between exposures received and the frequency and severity of adverse consequences, including infection, illness-days, and death. In terms of the causal chain:

act → exposures → illnesses ← covariates,

exposure-response modeling describes the "exposures → illnesses" link.

Multivariate dose-response models allow individual susceptibility and other covariates, as well as exposures, to be included explicitly in the calculation of adverse effect probabilities, i.e., they describe the conditional probability of adverse effects (both frequency and severity) for different exposures in different subpopulations, described by different combinations of age, sex, health status, or other covariates. Multivariate dose-response models quantify the following sub-diagram:

exposures → illnesses ← covariates

At present, univariate dose-response models are more commonly used in microbial risk assessment and antimicrobial risk assessment.

DESIRED OUTPUTS OF EXPOSURE-RESPONSE MODELING

A successful quantitative exposure-response model provides a mathematical function relating exposure levels (and possibly other covariates) to probabilities of adverse consequences. These probabilities may also be expressed in units of conditional expected frequencies, or rates per unit time, of specified adverse effects for exposed individuals. *Confidence bands* should be used to express uncertainties about consequence probabilities or rates at different exposure levels. Dose-response relations and their confidence bands may be presented for individuals in different identified subpopulations, e.g., those with special sensitivity or susceptibility, as well as for a randomly sampled individual.

Technical Note Classification trees for multivariate dose-response modeling: If it is desired to create a simple qualitative exposure-response model, then decision tree algorithms can be used to automatically bin more detailed exposure-response data into a few aggregate exposure intervals (or combinations of intervals, if there are multiple risk factors in a multivariate exposure-response model) that predict similar response levels. (See e.g., Zhang H, Singer B. *Recursive Partitioning in the Health Sciences.* New York: Springer; 1999.) Such approximate exposure-response relations can be smoothed in a number of ways to reconstruct simple, approximate exposure-response functions and confidence intervals directly from detailed exposure-response data (http://www.stat.wisc.edu/p/stat/ftp/pub/loh/treeprogs/guide/grapes.pdf). These techniques, while increasingly familiar and well-developed in biostatistics, are not yet commonly used in antimicrobial risk assessment.

EXAMPLES OF DOSE-RESPONSE MODELS

Appendix B presents several examples of dose-response models for *C. jejuni*. They differ primarily in their assumptions about inter-individual heterogeneity in individual dose-response parameters and in their assumptions about whether observed population illness rates are driven more by differences in individual *exposures* or by differences in individual *susceptibilities* to exposures. Based on these differences, the dose-response models considered (the Beta-Poisson, which allows for interindividual heterogeneity in response parameters, a log-exponential model suggested by CVM, and a linear no-threshold model) give very different predictions for the effects of a change in the frequency distribution of microbial loads per *C. jejuni*-contaminated serving of chicken consumed, leading to an 42-fold variation (from 0.3 to 13.9) for the dose-response ratio factor near the bottom of Table 4. In general, uncertainties about dose-response relations can often introduce several orders of magnitude of uncertainty into predictions about the probable human health effects of risk management interventions that affect microbial loads reaching consumers.

RISK CHARACTERIZATION

DEFINITION OF RISK CHARACTERIZATION

Risk characterization integrates hazard identification, exposure assessment, and dose-response information to determine the probable frequency and severity of adverse health effects that exposure to a hazard causes in a population. For example, the Joint FAO/WHO Expert Consultation defines risk characterization as the "integration of hazard identification, hazard characterization [i.e., dose-response or exposure-response relation] and exposure assessment into an estimation of the adverse effects likely to occur in a given population, including attendant uncertainties". The US FDA has used this definition in microbial risk assessment (http://www.foodsafety.gov/~dms/lmriskgl.html).

PURPOSES OF RISK CHARACTERIZATION

Risk characterization is intended to show the predicted probable frequency and severity of adverse human health consequences (and other adverse effects of concern) for different risk management decisions. It presents expected impacts and confidence intervals for the number and severity of adverse outcomes per capita and per unit time. Thus, *risk characterization relates decisions to their probable consequences in order to guide and inform improved risk-management decision-making.*

DESIRED OUTPUTS OF RISK CHARACTERIZATION

A successful risk characterization describes the spectrum of health outcomes and the occurrence of the microorganism and/or resistance determinants of concern (based on the hazard identification step); the frequency distribution of exposures in the population (with confidence limits) for different decisions; the confidence limits for the dose-response model; and the confidence limits for the predicted frequency and severity of adverse effects for different risk management decisions (Haas et al., 1999).

The outputs from risk characterization should include:

- *Expected risk metrics* (i.e., expected number of infections, illnesses of specified severity levels, mortalities, treatment failures, etc.) per year and in a lifetime for a randomly selected member of the population;
- *Confidence intervals* around the expected risk for a randomly selected individual;
- Expected risks and confidence intervals for members of identified *sensitive sub-populations* and *highly exposed sub-populations*;
- Expected numbers and confidence intervals for *total infections and illnesses with different levels of severity,* per year, per capita-year, and per capita-lifetime in the total population and in identified sub-populations.

These individual risk and population risk metrics should be provided for each risk management decision being considered.

CAVEATS ON RISK CHARACTERIZATION FOR RISK MANAGEMENT

Many antimicrobial risk assessments to date have ignored the human health risks that risk management intervention might create, focusing instead entirely only on the human health risks that they might reduce or prevent. This represents a breakdown in sound risk assessment and risk management, on a par with assessing financial risks of an investment or acquisition based on only one side of a balance sheet.

In general, rational risk management requires considering and comparing the *total* human health consequences, both favorable and adverse, of the risk management decision options being evaluated. Risk characterization owes to risk managers a complete accounting of the illnesses or adverse human health effects (and other adverse consequences of interest for decision-making) that a risk management intervention might *cause*, as well as of those that it might *prevent*. As illustrated in Table 4 and Table 5, the same basic format and logic (multiplicative modeling) can be used to do both.

Similarly, risk management decision processes that map perceptions or data about the current situation directly to recommended actions or interventions without first explicitly identifying the probable human health consequences of the recommended actions or comparing the probable consequences of alternative decision options are prescriptively unsound. They violate important normative principles of rational and effective decision-making, i.e., decision-making designed to bring about desired consequences. This failure to follow the requirements of consequence-driven decision making is a potentially important limitation of approaches such as CVM's Guidance # 152, which proceeds directly from judgments about risks (or about poor surrogates for components of risk, such as the judged importance of classes of compounds in human medicine, which presumably would not be changed by any decisions that might be taken) to recommended risk management decision options, but without identifying or comparing their probable human health consequences.

A superior approach, according to the most widely accepted principles of decision science (e.g., Cox, 2001, Chaptesr 5-7) is to use quantitative risk assessment information about the probable consequences of alternative interventions to eliminate dominated options and to choose the best from among those that remain.

EXAMPLES OF RISK CHARACTERIZATION USING UPPER BOUNDS

The Consequence portions of Tables 4 and 6 illustrate risk characterizations based on plausible upper bounds calculated within the RRRT framework. In both tables, the product of the preceding factors gives an estimate of the potentially preventable adverse human health consequences per year caused by current animal antibiotic use.

UNCERTAINTY AND SENSITIVITY ANALYSIS

DEFINITION OF UNCERTAINTY AND SENSITIVITY ANALYSIS

Uncertainty analysis describes both the *uncertainty* in estimated risks (e.g., the extent to which confidence intervals could potentially be narrowed by collecting more data) and also the *variability* in risk estimates, based on differences in individual exposure and response parameters. Uncertainty can be eliminated, at least in principle, by larger sample sizes and better information, whereas variability cannot be reduced no matter how much additional information is collected.

Sensitivity analysis identifies the inputs to a risk characterization that most affect the widths of confidence intervals for risks. Quantitative sensitivity analyses show how risk estimates (point estimates and confidence intervals) and recommended risk management decisions change as inputs are varied and as uncertainties in the input are reduced.

PURPOSES OF UNCERTAINTY AND SENSITIVITY ANALYSIS

Uncertainty and sensitivity analyses indicate the potential for conclusions about risks (including their uncertainties) to change as additional information is collected. The potential for change indicates how certain and robust are risk estimates and risk management decisions based on currently available information. Showing the potential for change if more information is obtained can provide affected stakeholders with a desirable incentive to collect additional relevant information in order to change current risk estimates and risk management decisions.

DESIRED OUTPUTS OF UNCERTAINTY AND SENSITIVITY ANALYSIS

Uncertainty analysis provides confidence intervals around risk estimates. Sensitivity analyses show how risk estimates and confidence interval widths are expected to vary as inputs are changed and as uncertainties (e.g., confidence interval widths) for inputs are reduced. Sensitivity analysis also shows how model results and predictions change if different plausible modeling assumptions are made. It is important to include the impacts of such modeling uncertainties, as well as of data uncertainties and sampling errors, in any complete uncertainty analysis. Technical methods for displaying the results of sensitivity analyses (e.g., tornado diagrams, spider diagrams) have been developed in the decision and risk analysis literatures and are appropriate to include in technical discussions and presentations of risk analysis results.

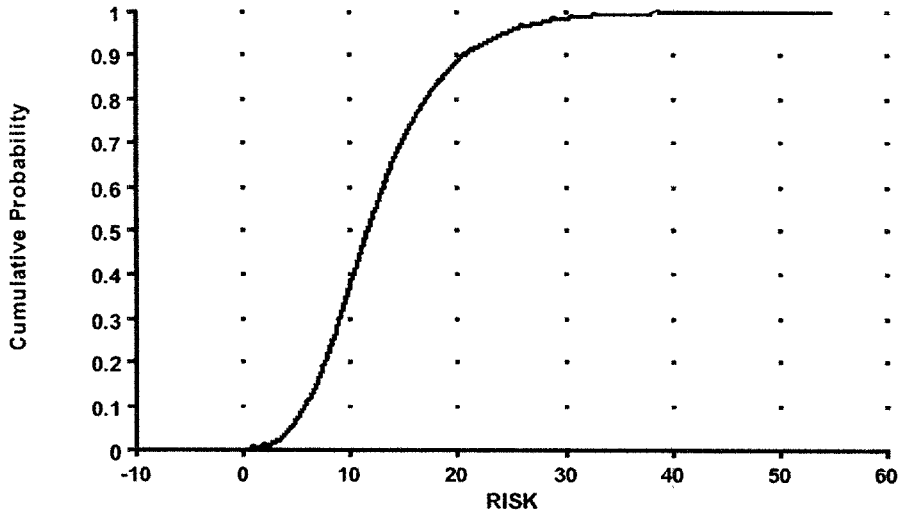
The output of a quantitative uncertainty analysis is a joint posterior probability distribution for model quantities and predictions after conditioning on observed data (and on modeling assumptions and uncertainties). These can be displayed as joint confidence regions for model parameters or predictions.

EXAMPLE OF AN UNCERTAINTY ANALYSIS OUTPUT

To illustrate one of the key concepts of uncertainty analysis, the following figure displays the uncertainty in a risk estimate calculated using the RRRT formula:

RISK = Exposure * Dose-Response * Consequence

where RISK is measured in expected number of excess illness-days per year, Exposure is measured in potentially infectious meals ingested per year in a population, Dose-Response = expected number of illnesses caused per potentially infectious meal ingested, and Consequence is measured in illness-days caused per illness. For purposes of illustration, the point estimates are taken to have median values of: Exposure = 100, Dose-Response = 0.2, and Consequence = 6 days. To express uncertainty, Exposure is modeled as a log-normal distributions with a geometric standard deviations of 1.4; Dose-Response is modeled as a Bernoulli random variable having value of 1 with probability 0.2 (for "susceptible" members of the exposed population) and a value of 0 otherwise (and hence a mean value of 0.2); and Consequence is modeled as a normally-distributed random variable with mean of 6 days and standard deviation of 2 days. The curve is a cumulative probability distribution for RISK, given these uncertain estimates of Exposure, Dose-Response, and Consequence. It was generated via using the Analytica™ Monte Carlo uncertainty analysis software (http://www.lumina.com/).



RISK MANAGEMENT

DEFINITION OF RISK MANAGEMENT

Formal risk management is a decision process that maps available information from risk assessment about the probable consequences of acts into choices of which acts to take. Acts available to risk management decision-makers usually include collecting additional information to reduce uncertainty about exposures and risks, as well as opportunities to disseminate existing information and warnings and to require or constrain activities by private agents.

Risk management decision models can be used to quantify the expected value of additional information for improving decision-making, and hence can help to set research priorities. They also prescribe interim decisions to be made unless and until additional information becomes available.

PURPOSES OF RISK MANAGEMENT

Risk management decision processes and institutions are used to prevent, mitigate, transfer, share, and spread risk and to assign liability and compensate victims of risks. Common options for risk management include the following:

- *Warn:* Inform or warn potential participants about risks of activities and transactions. For example, putting warning labels on food products may help consumers take sufficient care to avoid risky preparation or consumption practices.
- *Facilitate* voluntary risk management agreements by verifying and publicizing relevant risk information and specifying no-agreement outcomes.
- *Provide insurance:* Underwrite health care costs for specific illnesses to mitigate the financial component of losses.
- *Regulate:* Restrict voluntary activities or transactions (e.g., production, sale, or use of antimicrobial feed additives or other animal antibiotic uses) by imposing constraints, standards, and regulatory requirements based on risk information.
- *Use litigation and process design:* Design and enforce processes and rules (e.g., tort liability rules, inspection and labeling or licensing programs, workers' compensation)
- *Compensate:* Compensate known or suspected victims of hazardous activities, or force others (e.g., their known or suspected injurers, or tax payers) to pay compensation.

CAVEATS ON RISK CHARACTERIZATION FOR RISK MANAGEMENT

Because the following points are so important for many real-world applications, we repeat here the same warnings given at the end of the Risk Characterization section.

Many antimicrobial risk assessments to date have ignored the human health risks that risk management intervention might create, focusing instead entirely only on the human health risks that they might reduce or prevent. This represents a breakdown in sound risk assessment and risk management, on a par with assessing financial risks of an investment or acquisition based on only one side of a balance sheet.

In general, rational risk management requires considering and comparing the *total* human health consequences, both favorable and adverse, of the risk management decision options being evaluated. Risk characterization owes to risk managers a complete accounting of the illnesses or adverse human health effects (and other adverse consequences of interest for decision-making) that a risk management intervention might *cause*, as well as of those that it might *prevent*. As illustrated in Table 4 and Table 5, the same basic format and logic (multiplicative modeling) can be used to do both.

Similarly, risk management decision processes that map perceptions or data about the current situation directly to recommended actions or interventions without first explicitly identifying the probable human health consequences of the recommended actions or comparing the probable consequences of alternative decision options are prescriptively unsound. They violate important normative principles of rational and effective decision-making, i.e., decision-making designed to bring about desired consequences. This failure to follow the requirements of consequence-driven decision making is a potentially important limitation of approaches such as CVM's Guidance # 152, which proceeds directly from judgments about risks (or about poor surrogates for components of risk, such as the judged importance of classes of compounds in human medicine, which presumably would not be changed by any decisions that might be taken) to recommended risk management decision options, but without identifying or comparing their probable human health consequences.

A superior approach, according to the most widely accepted principles of decision science (see e.g., Cox, 2001, Chaptesr 5-7 for an in-depth review for risk analysts) is to use quantitative risk assessment information about the probable consequences of alternative interventions to eliminate dominated options and to choose the best from among those that remain.

RISK COMMUNICATION

DEFINITION OF RISK COMMUNICATION

Risk communication facilitates the effective participation and interaction of technical experts, stakeholders, and decision-makers in risk management decision processes and deliberations. Risk communication is also used to present the results of risk analyses to affected stakeholders, decision-makers, participants, and other audiences (http://www.foodsafetynetwork.ca/risk.htm#communication). Communication and deliberation drive much of the risk management decision process in many cultures and are essential for successful outcomes (http://www.belleonline.com/oct_02.pdf).

Communication of risk analysis results typically consists of the following steps: (a) Identify the goals and definitions of success for the risk communication effort (http://www.sirc.org/publik/revised_guidelines.shtml); (b) Select messages to be communicated to/discussed with each audience to achieve the goals; (c) Select framing, presentation media, displays, exhibits, interaction styles and formats, and a script for presenting the messages; (d) Implement the risk communication plan; and (e) Monitor results and incorporate feedback about the effects of the communication into a revised plan.

PURPOSES OF RISK COMMUNICATION

The most common goals for risk communication programs are *informing* individuals and groups about risks so that they can make better-informed decisions or seek more information; *influencing* people to change their behaviors, their attitudes and beliefs about hazards, and their acceptance of risk management decisions and policy recommendations; *involving* affected parties in the decision process; and *facilitating* their participation in conflict-resolution, consensus-building, and collective decision-making about risk management. The field of risk communication provides guidelines, derived mainly from experience, analysis of survey data, and experiments, for how to accomplish these goals by sharing risk information among stakeholders and decision-makers.

OIE (http://www.oie.int/eng/publicat/rt/2003/VOSE.PDF, *Rev sci. tech. OK int Epiz., 20 (3)*) states that "The goals of risk communication are the following:

- to promote awareness and understanding of the specific issues under consideration during the risk analysis process, by all participants
- to promote consistency and transparency in arriving at and implementing risk management decisions
- to provide a sound basis for understanding the risk management decisions proposed or implemented
- to improve the overall effectiveness and efficiency of the risk analysis process
- to strengthen working relationships and mutual respect among all participants

- to promote the appropriate involvement of all stakeholders in the risk communication process
- to exchange information on the knowledge, attitudes, values, practices and perceptions of stakeholders concerning the risks in question."

There is a tension in many risk communication efforts between *informing* and *influencing* or manipulating target audiences in presenting risk information (Ng and Hamby, 1997). Risk communication programs are often designed and evaluated based on their success in changing individual behaviors, e.g., by persuading people to stop eating fish from polluted lakes, to start using sun block, to participate in vaccination programs, to wear seatbelts, or to refrain from smoking. Other risk presentations have as their main goals to make decisions that have already been reached palatable to those affected (often a lost cause if those affected did not participate in the decision) and to confer legitimacy on decision processes by holding open meetings and sharing information.

Effective communication and facilitation about food-related risks enables stakeholders, experts, and decision-makers to participate more effectively in risk management decision processes. It does so by structuring how their beliefs, values, and concerns are elicited, shared, used to create and evaluate decision options, and acted on. It may also enable the facilitator to pursue policy goals in setting the agenda and managing the process to promote certain ends.

DESIRED OUTPUTS OF RISK COMMUNICATION

A successful risk communication program summarizes and presents the results of risk analysis in a way that clearly and credibly answers the following questions for the intended audience: (a) What should I do now? (b) Why is it desirable? / What are the benefits? (d) Why should I believe it?

The output of a risk communication program should be an exposition of risk analysis results that is both accurate and effective in changing beliefs, attitudes, and behaviors. Communication and presentation styles that are most effective in changing behaviors typically differ in structure, content, and emphasis from those that best express the technical content of risk assessment findings or that invite and elicit fruitful participation and interaction. For example, accurate communication of technical findings about risks and uncertainties to technically trained decision makers, and effective internal communication about facts, assumptions, conclusions, and uncertainties among expert members of a risk-assessment or risk management team, can greatly benefit from technical methods. Causal graph models, simulation-based what-if analyses, sensitivity analyses, risk profiles, and Bayesian posterior distributions (see Appendix A) can convey precisely what is known, how it is known, and what remains unknown or assumed – to audiences well trained in such methods.

But technically accurate risk communication does *not* address other key goals, such as telling people what has been decided or what they should do persuasively and credibly so that they agree (http://www.foodsafetynetwork.ca/food/blainepowell.pdf). It may not even give non-specialists the information they need to make improved decisions. It does not address the need to elicit stakeholder concerns and values or to address them in risk assessment and decision making. By contrast, persuasive communication about risks and risk management decisions to stakeholders, media, and the public requires different building trust, gaining and maintaining credibility, and preparing effective summaries of decision-relevant information using appropriate framing techniques. Brevity, clarity, focus, candor, cogent examples, and deliberate attempts to distance one's self from negative stereotypes of risk communicators may be crucial for communicating technical risks to non-specialist audiences so that the message is listened to instead of being tuned out or dismissed (Peters et al., 1997; Byrd and Cothern, 2000, Chapter 12.)

These factors help to establish an audience's perception of knowledge and expertise, openness and honesty, and concern and care – all of which, in turn, tend to promote trust in the speaker and acceptance of his or her risk messages. More generally, audience members consider the source of information, emotional style, framing, and imputed motives of the speaker in assessing the credibility of the message and in responding to it (http://www.inspection.gc.ca/english/corpaffr/publications/riscomm/riscomme.shtml).

METHODS FOR RISK COMMUNICATION

How risk information is formatted and presented can greatly affect how recipients assimilate and act on it. For example, in medical decisions, people are more likely to elect a medical procedure when it is described as "99% safe" then when it is described as having "1% chance of complications" (Gurm and Litaker, 2000). Presenting relative risks rather than absolute risks and using loss framing instead of gain framing make it more likely that patients will adopt screening procedures. In presenting chemical risks, the language used to describe risks may trigger speculations about the presenter's motives and undermine his or her credibility with the target audience (MacGregor et al., 1999). Understanding such effects can help in preparing the presentation of factual information in ways that are likely to elicit desired responses.

A striking insight from the framing literature is that *there may be no neutral way to present risk information.* Any presentation carries with it potential presentation and framing effects and biases that may affect the recipients' attention, interpretation, and actions. Presenting the same information in different ways and emphasizing fact-rich displays (e.g., cumulative risk profiles) that are not strongly associated with known presentation biases may come as close as possible to providing the information needed for rational decision-making without influencing the decision. Such displays often lack the brevity and focus that are most effective in action-oriented presentations.

Effective risk communication must be concerned with process as well as with outcome. If people believe that identifiable groups are having risks imposed on them unfairly by identified others having superior power, authority, or information, the result is likely to be outrage (Ng and Hamby, 1997). Unresolved outrage can quickly destroy the chances for joint problem-solving as an approach to risk management decision-making and conflict resolution. To resolve such situations, it is important to acknowledge and address the perceived unfair situation, either by correcting it or by discussing how decisions *should* be made when values and interests genuinely conflict and then demonstrating willingness to abide by agreed-to principles of fairness in deciding and communicating what will be done.

The following guidelines for communicating regulatory risk analyses and risk management decisions to the public are representative of much prescriptive literature on structuring risk communication and management efforts (e.g., Ng and Hamby, 1997).

Elements of a Successful Agency Risk Communication Plan

1. *Be clear on the roles and goals* of risk management program (e.g., is the goal to inform, influence, or involve the audience?)

2. *Address stakeholder concerns.* What knowledge, beliefs, values, attitudes, cultures, and contextual factors shape their concerns and motivate their actions?

3. *Study/understand risk perceptions*, concerns, and most effective communication styles.

4. *Involve stakeholders.* Successful risk communication should be interactive and participatory, not a one-way broadcast.

5. Develop technical risk assessment content to support effective risk communication by answering specific questions/addressing concerns. Emphasize decisions and consequences, not pure science

6. *Organize risk assessments to facilitate effective presentation* of content. Identify outcomes of interest or concern to stakeholders, identify decision options, show how they affect outcome probabilities, and quantify trade-offs among likely consequences of different options.

7. *Organize risk management decision processes* to eliminate outrage, accomplish goals, serve chosen roles, and reflect Agency values.

APPENDICES

APPENDIX A: METHODS AND DATA SOURCES FOR RISK ASSESSMENT

The achieve the goals and to produce the desired outputs of risk analysis for food-borne pathogens (resistant or not), it is necessary to have access to appropriate technical methods and data. This appendix introduces methods, references, and data sources useful in carrying out risk analyses. It briefly addresses the issues of how to create predictive risk assessment models, how to validate them, and how to work with realistic remaining uncertainties and data gaps. For additional discussion of methods and data, see Haas et al. (1999) (http://www.bookhq.com/compare/0471183970.html).

METHODS AND DATA FOR RISK ANALYSIS

Overall approaches for carrying out food safety assessments, microbial risk assessments, and antimicrobial risk assessments include:

- *Qualitative screening and ranking methods.* These seek to simplify and clarify risk analysis processes by using holistic scores or rankings for various components that contribute to exposure and risk (e.g., release of resistant strains, exposure to released strains, and consequences of exposure for reduced treatment effectiveness and/or for increased mortality, morbidity, or virulence.) These methods are still under development and require validation before their performance in identifying effective risk management decisions will be understood. Strengths include simplicity; potential limitations include unclear interpretations, inadequate information output to support rational risk management decision-making, and incorrect ranking of risks and/or of risk management interventions.

- *Quantitative risk analysis methods.* These seek to quantify how exposures and resulting frequencies, magnitudes, and population distributions of specific adverse health consequences will change if different risk management decisions are made. Examples of quantitative risk assessment approaches include "bottom-up" farm-to-fork risk models that seek to simulate or quantify how microbial loads change at each stage (e.g., transportation, slaughter, wholesale storage, retail, preparation); and "top down" retrospective ("fault tree") models that estimate the fraction of adverse health outcomes (such as treatment failures) that are caused by bacteria from specific sources. Strengths include providing information that is directly relevant for risk management, as well as ability to express both knowledge and uncertainty in considerable detail (e.g., using probability distributions.) Limitations include the need to collect relevant data to relate decisions to their probable impacts on exposures, as well as a need for data on exposure-response relations.

- *Decision trees.* Decision trees ask a sequence of qualitative and/or quantitative questions about a situation, then provide a risk estimate and/or recommended risk management action for the specific situation based on the answers to these questions (i.e., on which "tip" of the tree the answers lead to). The risk estimates at the tips of a tree may be quantitative or qualitative. Thus, decision trees can

combine aspects of qualitative and quantitative risk management. Strengths include simplicity, flexibility, and ability to incorporate all relevant factors in making risk predictions and recommendations. The main limitation is that all problems leading to the same tip of the tree are treated the same.

No matter which of these high-level approaches is used for organizing and displaying the results of the analysis, the process of risk analysis is traditionally presented as a sequence of steps, as in Table 1 of the text. The steps of hazard identification, dose-response modeling, and exposure assessment, typically require the specialized substantive knowledge of engineers, epidemiologists, industrial hygienists, toxicologists, and other scientists and subject-matter experts. Risk quantification and characterization use statistical, probability, process engineering, and simulation modeling methods to prepare probabilistic summaries of the likely frequency and severity (and perhaps also the distribution in a population) of adverse health consequences from different risk management alternatives. Back-tracking and iterative refinements at each step in Table 1 may occur as more information is gained. Quantitative risk characterization draws together all of the technical information, synthesizing hazard, exposure, and quantitative dose-response information into summaries of risk and uncertainty to inform risk management decisions. These decisions, in turn, may generate actions and information that lead to new information about hazards, exposures, and dose-response relations. It is now widely accepted that uncertainties about these components of risk must be included in the risk characterization.

Most health risks are highly uncertain. Recommendations to begin expensive risk management activities commonly trigger re-analyses of hazards, exposures, and plausible risk magnitudes as those who must bear the costs of risk management question the need for and health benefits from their efforts. Each round of risk characterization and risk management decisions can prompt new research into underlying causal mechanisms of effects and details of exposure. The results may eventually lead to revised risk estimates and recommendations for risk management. Thus, in practice, the steps in the risk analysis process are tightly linked and improvements in information at one step may trigger further iterations elsewhere in the process.

ISSUES AND CHALLENGES FOR RISK ANALYSIS

Key issues in performing a health risk analysis include:

- *Causal modeling of exposures and health effects.* In health risk analysis, opinions and beliefs about the probable consequences of decisions are expressed as explicit, publicly disclosed quantitative models of exposure and of exposure-response relations. These models usually represent aspects of *causation.* They are intended to describe how changes in actions propagate through one or more causal chains to change exposures and health effects. This emphasis on causation contrasts with many statistical and biostatistical models that emphasize *inference* about the probability of observing an effect, given observed data about an exposure. Statistical associations useful for

inference are not necessarily useful for predicting the causal impacts of changes in exposures.

- *Inferred risks and consequences.* Risks often cannot be directly observed or measured. Available epidemiological data may be insufficient to uniquely identify the contribution to health risks from specific microbial sources and resistance determinants. Therefore, specific risks must often be inferred from models, knowledge, assumptions, and data. Inferences about human health risks drawn from animal studies and/or from chemical structures and experimental results in various in vitro assays and biological systems require drawing conclusions based on analogy, induction from examples, and perhaps deduction about likely biochemical processes – forms of inference that are often challenging compared to routine statistical calculations of conditional probabilities or expected values.

- *Multiple stakeholders and decision makers.* Multiple stakeholders and decision makers usually participate in activities that create health risks and in risk management decision-making. For example, consumer exposure to a food-borne pathogen arises from the joint decisions and behaviors of farmers, producers, and food handlers and preparers. Multiple parties are therefore involved in risk management, and effective risk management may require coordinating their roles and responsibilities.

- *Risk management evaluation of health outcomes.* Evaluating risks that involve potential losses of life and health requires special techniques. Asking people about their willingness to pay (WTP) to reduce or remove such risks, or about their willingness to accept (WTA) monetary compensation to bear them, elicits responses that may reflect political and psychological attitudes and beliefs about rights and concerns about fairness and equity in the allocation of risks or costs while being quite insensitive to the probabilities and magnitudes of the consequences involved (Kahneman et al., 1999). Thus, individual economic values and preference trade-off rates for risks involving potential loss of health or life may be difficult or impossible to define and measure apart from political concerns and cultural attitudes.

- *Risk communication:* Risks and uncertainties must often be communicated to stakeholders who lack the data, resources, or expertise to accurately quantify and understand risks without assistance. Yet, objective communication of risk information is often difficult or impossible. How to communicate risk information without misleading or manipulating the audience is a substantial challenge.

- *Risk perception, comprehension, internalization* and *action:* Understanding and perceptions of risks by producers, retailers, consumers, and physicians, and their resulting behaviors, may be affected in unexpected ways by attempts to describe the results of formal risk assessments. Even if risk information is communicated accurately, behaviors may not change in the ways that models of rational behavior predict. Understanding how stakeholders (e.g., consumers, producers, physicians, and regulators) do and should make decisions based upon risk information is an active area of research.

METHODS OF RISK ANALYSIS THAT SHOULD BE AVOIDED

As the relatively new field of antimicrobial risk assessment has developed, several types of risk analysis have been attempted that, we believe, will ultimately be replaced by more useful and valid methods. These include:

- *Subjective/judgmental risk analysis.* While subjective rating and ranking systems can be useful for preliminary screening of hazards and risk management actions, they have important limitations. Subjective risk analyses in which participants are asked or required to make holistic judgments of risks attributable to a cause, or of the relative frequency and/or importance of hazards, exposures, or health consequences, are generally not trustworthy. They can lead to incorrect or ineffective risk management decisions (i.e., decisions that do not increase the probabilities of preferred outcomes), encourage group-think (including agreement and high perceived confidence in mistaken judgments and conclusions), and obscure or replace the rational and scientific reasons for decisions. Even technically trained experts are notoriously poor at forming accurate judgments about probabilities, causes, and risks (Kahneman D, Slovic P and Tversky A 1982. *Judgement Under Uncertainty: Heuristics and biases.* Cambridge University Press.) Subjective risk analysis should not be used as the primary basis for final risk management decision-making.

- *Data-free/Assumption-Driven Risk Modeling.* Some risk assessments in microbial safety and in antimicrobial risk analysis have created detailed sub-models (e.g., of cooking processes and their effects on microbial loads) based primarily or entirely on modeling assumptions, in the absence of adequate data to validate the modeling. Such assumption-driven modeling should be avoided: only models and assumptions that have been validated with data should be used for risk analysis. [In general, data-free modeling can be avoided by conditioning risk (i.e., probability and severity of adverse health effects) only on the observed causal predecessors (i.e., exposure and covariate information) of adverse health effects. Conditioning on all and only the available data leads to informed risk analyses while avoiding drawing conclusions that are highly dependent on untested, possibly incorrect, assumptions about unmeasured quantities.]

- Similarly, risk analysis based on *default assumptions,* or on *assumed attributions* of risks to specific causes, should be avoided unless and until the assumptions have been critically assessed and validated using data for the specific microorganisms, exposure routes, and populations for which risks are being evaluated. Specific combinations of microbial pathogens, transmission pathways (including food animal species, other foods, drinking water, swimming water, contact with animals, contact with infected people, etc.) should be assessed individually.

VALIDATING RISK ANALYSIS RESULTS

A risk assessment model predicts the probable human health effects and other consequences of different risk management decisions by predicting their impacts on human exposures to microbial loads. Following implementation of a risk management decision, these predictions should be tested by conducting an *evaluation study* to assess whether the predicted changes in exposures and health effects actually occurred. If not, the risk assessment may need to be refined (see Validation of Risk Characterization) and the recommended risk management decision may have to be revised.

METHODS AND DATA FOR HAZARD IDENTIFICATION

On the biological side, hazard identification draws on knowledge of infectious diseases, epidemiological data, and clinical microbiology to create testable hypotheses about causal relations among decisions, exposures, and health consequences (Haas CN, Rose JB, Gerba CP. *Quantitative Microbial Risk Assessment.* Wiley, 1999). On the statistical side, hazard identification uses methods of causal analysis to identify the decision-exposure and the exposure-health effects links in the following causal chain:

decisions → exposures → health effects ← covariates.

[In such a diagram, the conditional probability distribution of each quantity depends on the values of the quantities that point into it, if any. For details, see Chapter 1 of Shipley, B. Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference. Cambridge University Press. 2002. A link in a causal chain or graph is *identified* by using appropriate statistical tests to show that it is conditionally independent of all of its more remote ancestors, given the values of the variables that directly point into it. The causal relations are *quantified* by specifying the conditional probability of each variable's value, given the values of variables that point into it, if any. Input variables having no arrows pointing into them may have (unconditional) probability distributions reflecting uncertainty about their values.] Hazard identification deals with the identification of causal links that are then quantified in the exposure assessment and exposure-response steps.

Table A1 outlines steps for forming and testing causal hypotheses about exposure-response relation using epidemiological data. The more of these steps can be completed, the stronger is the inference that there is a causal relation between exposure and risk. Most statistical methods used in epidemiological risk analysis focus on steps 1 to 3, i.e., identifying non-random associations and eliminating potential biases and confounders as likely explanations. These steps can often be carried out using data from observational studies without requiring direct manipulations and experimental verification of predictions. The main method is to systematically enumerate and then eliminate (if possible) competing, non-causal explanations for the observed data using statistical tests.

Table A1: Steps to Establish a Causal Exposure-Risk Relation

1. *Identify a statistically significant exposure-response association.* Demonstrate that there is a non-random positive statistical association between exposure histories or events and adverse human health consequences (or other undesired consequences) in an epidemiological data set. Case-control, prospective cohort, or other cross-sectional or longitudinal epidemiological data may be used for this purpose.

2. *Eliminate confounding* as a possible cause of the association. Show that it is not due to or explained by other (non-exposure) causes such as differences in lifestyle factors, age, or exposures to other confounders. (Nurminen, 1983; Lin et al., 1998)

3. *Eliminate biases from sampling, information collection, and modeling choices as possible causes.* Show that the association is not explained by biases in who was selected (as study subjects or as controls) or in how information about them was collected and analyzed. (Choi and Noseworthy, 1992)

4. *Test and confirm hypothesized causal ordering and conditional independence* relations among observed values of variables. For example, show that the response is not conditionally independent of its hypothesized direct causal predecessors (e.g., exposure), but that it is conditionally independent of more remote causal predecessors given the direct predecessors. (Shipley, 2000)

5. *Confirm efficacy of interventions.* Confirm that changes in the levels of direct causal predecessors (e.g., exposures) are followed by the predicted changes in the levels of the variables they affect. This may often be done from time series observations, even if direct experimental manipulation is impossible using methods for interrupted time series analysis, intervention analysis, and quasi-experiment design and inference (Granger, 1980; Campbell and Stanley, 1963; McDowall et al., 1980).

6. *Identify and elucidate causal mechanism(s).* Explain how changes propagate via one or more causal paths to produce effects (Nurminen, 1997). A "causal path" is a sequence of steps in which completion of the earlier steps creates conditions that trigger or increase occurrence rates of subsequent steps. Such steps may be identified from experimental data and/or by applying generally accepted laws (Renton, 1994).

Many epidemiologists have recognized the logical necessity, in order to draw valid causal inferences, of refuting competing hypothesized explanations for observed exposure-response associations (Maclure, 1991). Table A2 summarizes common competing explanations (mainly, confounding and/or sampling, information, or modeling biases) and technical methods that have been developed to refute them. Yet, requiring alternative (non-causal) explanations to be refuted can be perceived as unhelpful when it is confined to merely identifying logical possibilities without also addressing their plausibility and the likely magnitudes of their impacts on risk estimates. For example, Savitz et al. (1990) state that "Biases that challenge a causal interpretation can always be hypothesized. ...It is essential to go beyond enumerating scenarios of bias by clearly distinguishing the improbable from the probable and the important from the unimportant." They argue that those who do not like a causal interpretation of epidemiological data might readily construct speculative hypothetical potential biases and confounders that can not all be refuted with available resources. This strategy could prevent conclusions about causation from being drawn when common sense and sound policy would be better served by accepting that causation is plausible, even if it is not practical to refute all conceivable alternative explanations.

On the other hand, accepting a statistical association as causal without rigorously examining and excluding competing hypotheses may make it too easy to launch expensive risk management control actions that would be effective if the association were causal, but that will not produce the anticipated benefits in practice. A partial solution to this dilemma is to focus on those non-causal explanations that appear to be *likely* and *important* (Savitz et al., 1990) i.e., those (if any) that might plausibly explain most or all of the observed exposure-response associations. Appropriate data analysis methods can often reveal which potential biases and confounders are most likely to provide non-causal explanations in specific studies. They can also help to eliminate logically conceivable biases that do not, in practice, play a large role. Most importantly, they can help to eliminate the most likely and important non-causal explanations when those do not, in fact, apply. Evidence that makes non-causal explanations unlikely makes causal explanations more likely, even if the evidence is not definitive.

In summary, the refutation approach to hazard identification suggests a key criterion for establishing causation for an observed exposure-response association: *have competing non-causal explanations been eliminated?* If so, then the hypothesis of causation is supported by the data used to refute them.

METHODS OF HAZARD IDENTIFICATION THAT SHOULD BE AVOIDED

Several common errors in hazard identification should be avoided, as follows:

- *Non-causal hazard identification.* Simply identifying a microorganism in food does not show that it causes human health harm in practice and thus does not constitute hazard identification. Even if it is known that the identified agent can be pathogenic under laboratory conditions, it is still necessary for hazard identification to demonstrate that it causes harm in reality.

Table A2: Potential Non-Causal Explanations and Refutations for Exposure-Response Associations

Potential Non-Causal Explanation	Methods to Avoid or Refute
Modeling Biases	
Variable selection bias (including selection of covariates to include in the model)	Bootstrap variable selection, Bayesian model averaging (BMA), cross-validation for variable selection.
Omitted explanatory variables (including omitted confounders and/or risk factors)	Include potential confounders in an explicit causal graph model; test for unobserved latent variables,
Variable coding bias (i.e., how variables are coded may affect apparent risks)	Use automated variable-coding methods (e.g., classification trees). Don't code/discretize continuous variables.
Aggregation bias / Simpson's paradox	Test hypothesized relations at multiple levels of aggregation. Include potential confounders in an explicit causal graph model.
Multiple testing/multiple comparisons bias	Use current (step-down) procedures to adjust p-values without sacrificing power.
Choice of exposure, dose, and response metrics	Use multiple exposure indicators (e.g., concentration and time). (Don't combine.) Define responses as survival functions and/or transition rates among observed health states.
Model form selection bias and uncertainty about the correct model for exposure-response relation and other relations.	Use flexible non-parametric models (e.g., smoothers, wavelets); Bayesian Model-Averaging. Report model diagnostics and sensitivity analyses of results to model forms
Missing data values can bias results	Use data augmentation, EM algorithm, MCMC algorithms
Measurement and misclassification errors in explanatory variables	Use Bayesian measurement error models, data augmentation, EM algorithm, and other missing-data techniques
Omitted heterogeneity in individual response probabilities/parameters	Latent variable and mixture distribution models, frailty models of inter-individual variability
Bises in interpreting and reporting results	Report results (e.g., full posterior PDFs) *conditioned* on choices of data, models, assumptions, and statistical methods. Show sensitivities of results to these choices.
Sample Selection Biases	
Sample selection (sample does not represent population for which inferences are drawn)	Randomly sample *all* cohort members if possible
Data set selection bias (i.e., selection of a subset of available studies may affect results)	Use meta-analysis to show sensitivity of conclusions to studies. Use causal graph models to integrate diverse data sets
Health status confounding/Hospital admission bias (as well as referral and exclusion biases)	• If possible, use prospective cohort design (Feinstein, 1988) • "Use population-based cases and population-based controls" (Choi and Noseworthy, 1992)
Selective attrition/survival (e.g., if exposure affects attrition rates) Differential follow-up loss	• Use a well-specified cohort (Feinstein, 1988) • "Include non-surviving subjects in the study through proxy interviews" (Choi and Noseworthy, 1992) • Compare counter-factual survival curves (Robins, 1987)
Detection/surveillance bias	• Match cases to controls (or exposed to unexposed subjects) based on cause of admission.
Membership bias (e.g., lifestyle bias, socioeconomic history)	• In cohort studies, use multiple comparison cohorts. • Hard to control in case-control studies.
Self-selection bias; Response/volunteer bias	• Achieve response rate of at least 80% by repeated efforts • Compare respondents with sample of non-respondents
Information Collection Biases	
Intra-interviewer bias	• Blind interviewers to study hypotheses, subject classifications
Inter-interviewer bias	• Use same interviewer for study and comparison groups
Questionnaire bias	• Use dummy questions to mask study goals • Avoid leading questions/ leading response options
Diagnostic suspicion bias Exposure suspicion bias	• Hard to prevent in case-control studies. In cohort studies, make diagnosis and exposure assessments blind to each other.
Recall bias • Differential • Non-differential	• Conduct sensitivity analysis of recall bias during analysis. (Barry, 1996; Greenland, 1996) • Use exposure proxies to check for recall bias. (Basso, 1997) • Apply bias-correction algorithms

- *Hazard identification based on temporal trends.* Discussions of hazard identification in antimicrobial risk assessment sometimes refer to temporal trends, in which adverse health effects occurred after the historical introduction of an antimicrobial agent in feeds. Such discussions are usually inconclusive, neither establishing nor refuting the possibility of a causal relation. In rigorous analyses of causation, such "temporal trend" arguments are dismissed as instances of the *ex post* or "false cause" fallacy (e.g., http://library.stmarytx.edu/lac/IStudies/fallacyG.htm; http://www.beige.org/~gltweasl/fallacy.htm.) However, if sufficient longitudinal data are available, then they may be used in conjunction with correct definitions of causation (e.g., http://instruct.uwo.ca/fim-lis/504/504ter.htm) and objective tests and analytic methods for causation (e.g., http://citeseer.nj.nec.com/55450.html) for purposes of hazard identification and risk assessment. However, the relevant methods (Table A1) are not based simply on trends, but on showing that significant changes in time series occur following an intervention, such as introduction of a feed additive. Statistical methods of intervention analysis and change point analysis in time series can be used for this purpose.

- *Unspecified harm.* Identifying a pathogen without identifying any clinically relevant harm that it causes does not constitute adequate hazard identification. A successful hazard identification must demonstrate that exposure to the identified pathogen actually causes a specific adverse health effect, rather than merely that the potential causal agent is present. *Although purely hypothetical or theoretical suppositions about harm caused by microorganisms or resistance determinants in food can be useful as hypothesis-generating steps, they do not constitute hazard identification.* Instead, hazard identification requires that the causal hypotheses be tested objectively, and that competing explanations be identified and refuted, using appropriate statistical methods.

- *Partial or incomplete hazard identification.* Identifying only one component of hazard (e.g., effects of animal antimicrobial agents on human health effects of resistant bacteria but not susceptible bacteria) can give an incomplete description of potential risk that is not suitable for guiding rational decision-making.

In summary, although it may be difficult or impossible to prove causation from data, it usually *is* possible to test and refute competing non-causal hypotheses and to test whether the hypothesized causal relations between decisions and exposure and between exposure and risk are consistent with available data. These tests, and the presentation of the results, are the essential outputs of a successful hazard identification.

METHODS AND DATA FOR EXPOSURE ASSESSMENT

Exposure assessment uses predictive microbiology to predict how microbial loads reaching consumers or other exposed populations (e.g., patients) will change if different risk management decisions are made. The required data and calculations can be organized and presented using any of the following exposure modeling approaches:

- *Process simulation modeling approaches* (Haas et al., 1999, 225-248) describe the flow of food animal carcasses, portions, and products through various sub-processes, each characterized by an input-output relation described by a regression model or other simple statistical or simulation model. Changes in microbial loads are tracked as part of the input-output descriptions. Available measurements and data may be used to fit simple probability distributions and parametric models to describe the growth or attenuation of transmitted microbial load at each stage. Examples include Poisson or negative binomial distributions of microbial loads (fit using most probable number (MPN) data and maximum-likelihood statistical estimation algorithms) and Gompertz growth curves for pathogen growth kinetics. Data for growth rates of *E. coli* O157:H7, *Listeria monocytogenes, Clostridium botulinum, Staphylococcus aureus,* and other common pathogens can be found in Haas et al. (1999, Chapter 6) and its references. Consumption factors and frequencies for water and foods (beef, fish, chickens, eggs, shellfish, etc.) are available from the literature (*ibid,* p. 239, 241) and can be used to model the frequency with which microbial loads on food portions are ingested. Transfer rates of bacteria between skin and hands and from food to hands have also been estimated for various bacteria; however such details are not necessarily needed or useful if adequate data on earlier and later points in the causal chain leading from animal loads to human illness are available.

- *Farm-to-fork models* are an important type of process simulation model. Farm-to-fork models track the microbial load distributions on animal carcasses, portions, and servings through successive stages of production, processing, transport, slaughter storage, preparation, and consumption. Monte Carlo simulations of the probabilistic input-output relations at each stage are used to propagate microbial load frequency distributions throughout the model. A closely related methodology is *dynamic flow tree modeling* for microbial risk assessment (Marks HM, Coleman ME, Lin CT, Roberts T. Topics in microbial risk assessment: dynamic flow tree process. Risk Analysis 1998 Jun;18(3):309-28; http://www.nap.edu/books/0309086272/html/97.html). Dynamic flow trees use Monte Carlo uncertainty analysis to sum risks over many scenarios, weighted by their respective expected frequencies or probabilities, but without necessarily tracking all the process steps in a farm-to-form model.

- *Retrospective attribution models* begin with adverse health outcomes and/or clinical measurements of bacteria-related harm. They use genotyping and other data to estimate the fraction of these cases of adverse effects that could have been caused

by specific exposures and sources. This attribution of measured outcomes to the sources that contribute to them may be repeated in several stages. Thus, the calculation works backward from clinical outcomes to intermediate sources (e.g., community vs. nosocomially acquired sources), and then to their predecessors (e.g., food or water consumption, contact with infected or contaminated animals and humans, etc.), and so forth. The process is repeated, ultimately leading back to a fraction of microbial load (resistant, susceptible, or both) that is estimated to be due to use of antimicrobial agents on the farm. At each stage, the fraction of outcomes contributed by each preceding source is estimated. The fractions along the path leading from adverse outcomes back to antimicrobial use on the farm are then multiplied and the results are interpreted as estimates of the fraction of cases per year that could be prevented by eliminating any of the steps along the path, including antimicrobial use on the farm.

In general, exposure models describe the dispersion and transport of hazardous materials (e.g., concentrations of CFUs on food animal products) through different media and pathways (e.g., food, water) leading from the source(s) to members of the exposed population. Process sub-models link the source strength and the positions of target receptors, perhaps integrated over time, to predict quantitative exposures received by the targets. In addition, exposure models model the distribution over time of human populations among locations (e.g., restaurants, kitchens) and activities that result in exposures.

DEALING WITH UNCERTAIN AND INCOMPLETE EXPOSURE DATA

A common misunderstanding about quantitative exposure models is the mistaken idea that they require unrealistic amounts of detailed data. This is not true when appropriate techniques of uncertainty analysis are used. For example, suppose that there are several consecutive stages in a process simulation exposure model, such as:

$$A \Rightarrow B \Rightarrow C \Rightarrow D$$

Here, the wide arrows represent input-output processes and A, B, C, D are points where microbial loads might be measured. If measurements are unavailable for stage C, then the probability distribution of microbial loads at D can still be related to microbial loads at A (thus leaving the chain unbroken by the missing data at C) via the formula:

$$\Pr(\text{load at } D = d \mid \text{load at } A = a) = \Sigma_b \Pr(\text{load at } B = b \mid \text{load at } A = a)\Pr(\text{load at } D = d \mid \text{load at } B = b)$$

where

$$\Pr(\text{load at } D = d \mid \text{load at } B = b) = \Sigma_c \Pr(\text{load at } D = d \mid \text{load at } C = c)\Pr(\text{load at } C = c \mid \text{load at } B = b)$$

In other words, it is possible to *condition on what is observed* while skipping over (or "marginalizing out", in statistical terminology) unobserved quantities by summing over all

possible values, weighted by their conditional probabilities. The basis for inferring exposure distributions is conditioning of probability distributions for output quantities within a causal model (the simulation model), not exhaustive simulation of all relevant details. Statistical inference algorithms (e.g., the data augmentation algorithm; see Schafer JL. *Analysis of Incomplete Multivariate Data.* New York: Chapman and Hall;1997) allow the conditional distributions of outputs to be quantified, conditioned on observed data, even when other data are missing. Thus, there is much flexibility within the process simulation approaches to use all available data without requiring use of unavailable data.

In retrospective attribution exposure models, uncertainties about the attributable fractions at different stages are commonly treated by using upper-bound estimates. If the product of the upper bound estimates is small, then the true but unknown value of the product is also small, and this information may be sufficient to support a decision that no intervention is required. If the product of the upper bound estimates is large enough so that this conclusion cannot be justified, then the uncertainty analysis can be refined by estimating probability distributions for the fractions at different stages and applying Monte Carlo uncertainty analysis to obtain the probability distribution for their product.

METHODS OF EXPOSURE ASSESSMENT TO AVOID

The following approaches to exposure assessment seek to simplify the process by eliminating the need to consider microbial load distributions. In general, they give inaccurate results and should be avoided.

- *Holistic statistical models.* An example is the simple linear regression model:

  Exposure received = k*contamination at farm

where k is interpreted as an overall transmission coefficient. The problems with such a model (called a "reduced form model") are that (a) It fails to sum over multiple paths and scenarios (which must typically be represented by multiple distinct k values); and (b) It does not necessarily represent information about how a change in the right-side explanatory variables ("contamination at farm", in this example) would causally affect the left-side variable.

Note: Only structural equations, not reduced-form ones, reveal the causal relations among variables (Shipley, 2000). As an extreme hypothetical example to illustrate this point, suppose that the correct structural equations in a model are:

Exposure received = contamination at retail – contamination in kitchen (1)

contamination in kitchen = (1/3)*contamination at retail (2).

Equation (2) is mathematically, though not causally, equivalent to:

contamination at retail = 3*contamination in kitchen (2').

Substituting (2') into (1) gives the reduced-form model

Exposure received = 2*contamination in kitchen. (3)

While this is a valid equation for statistical inference, it cannot be used to correctly predict the causal effect on "Exposure received" of increasing "contamination in kitchen". Equation (1) shows that this effect is negative (i.e., each unit increase in "contamination in kitchen" decreases "Exposure received" by one unit), while equation (3) indicates a positive statistical relation between them. Equation (1), rather than equation (3), is relevant for predicting causal impacts of interventions.

- *Prevalence-based exposure metrics.* In general, models that use binary summaries of exposure metrics (i.e., "contaminated" vs. "not contaminated") do not contain enough relevant exposure information to make accurate risk predictions (e.g., Rosenquist et al., 2002). Such prevalence summaries of exposure should not be used for risk assessment.

Note: If a qualitative summary of exposures is desired, then instead of using prevalence metrics, decision tree algorithms can be used to automatically bin more detailed exposure-response data into a few aggregate exposure intervals (or combinations of intervals, if there are multiple exposure factors) that predict similar response levels (Zhang H, Singer B. *Recursive Partitioning in the Health Sciences.* New York:Springer; 1999.)

VALIDATION AND REFINEMENT OF EXPOSURE ASSESSMENT MODELS

Exposure assessment models should be validated by comparing their predictions under different conditions to measured values of exposures and/or their surrogates. An exposure model is used to predict microbial loads (or their surrogates) at measurement points under different conditions, e.g., in different locations, for different seasons, etc. These predictions are compared to the measured values using statistical goodness-of-fit tests and diagnostic plots, to determine whether the observed values are statistically significantly different from the predicted distributions of values. Haas et al. (1999), Chapter 6, provides details of goodness-of-fit tests for parametric exposure models, (Chapter 7 also discusses validation and uncertainty analysis of simulation models, although with emphasis on dose-response rather than exposure models.)

If the predicted exposures do not adequately match validation data, then the exposure model should be corrected. This can be done by refining the model to include omitted variables, to more accurately model dependencies among its inputs (Haas CN., On modeling correlated random variables in risk assessment, Risk Anal 1999 Dec;19(6):1205-14) and/or by using the differences between predicted and observed values to select more appropriate mathematical model forms that can explain and reduce these differences. To avoid using non-valid models, it is often desirable to apply flexible non-parametric data descriptions and to use predictions from multiple alternative models that are consistent with

available knowledge and data, weighted by their probabilities. Table A2 mentions technical methods and statistical algorithms for accomplishing these goals.

If the predicted exposures adequately match validation data according to goodness-of-fit tests and model diagnostics (e.g., plots of residuals), then the exposure model may be used to make predictions for risk assessment within the validated range of conditions. In this case, remaining uncertainty in model parameters, inputs. and predictions should be expressed through *confidence intervals* for single quantities (e.g., the mean exposure or the upper 95% exposure limit in the population) and through *joint confidence regions* for multiple quantities, such as the exposures received by different subpopulations.

METHODS AND DATA FOR DOSE-RESPONSE MODELING

The main approaches to quantitative exposure-response modeling are as follows:

- *Empirical parametric statistical dose-response models:* If experimental data are available, e.g., from feeding studies, then any parametric statistical model that adequately describes the dose-response data may be used to summarize it (at least over the range of the observed data.) For example, logit, probit, log-logistic, log-probit, and Weibull models are often used. These empirical model curves can be fit automatically to data using maximum-likelihood estimation (MLE) or other statistical curve-fitting algorithms. However, extrapolating empirical statistical models outside the range of the experimental data, especially, to low doses, is in general not justified. Empirical curve-fitting should be viewed as an interpolation approach for describing potentially large amounts of experimental data with a smaller number of parameters.

- *Biologically motivated parametric dose-response models.* The probability that enough ingested organisms survive to reach a site where infection can be initiated has been calculated in simplified biomathematical models of the probabilistic survival and infection processes. Doing so leads to a catalog of parametric dose-response models appropriate for different simplifying assumptions about the disease process. These include the exponential, one-hit, multi-hit, Beta-Poisson, threshold, and negative binomial models, as well as mixture distribution models for populations with heterogeneous dose-response parameters. These parametric models can be fit to experimental data using maximum-likelihood estimation (MLE) or other statistical curve-fitting algorithms. Unlike the purely empirical models, these models provide a theoretical basis for extrapolating beyond the range of the data used to fit them, at least to the extent that their underlying assumptions provide useful approximate descriptions of biological reality.

- *Complex dose-response models.* Dose-response relations can be decomposed into components, e.g., modeling the internal dose received from a given external dose applied; the probability of infection given dose; and the probability of illness given infection. If these components can be estimated separately from available data, then the results can be composed to give an estimate for the total exposure-illness dose response function. Moreover, separate estimation of components may provide a way to help extrapolate results across species, if similarities and differences in relevant component processes are known. In practice, however, this decomposition strategy has usually been combined with simplifying assumptions about the components, leading to the biologically motivated parametric dose-response models already mentioned.

- *Epidemiological exposure-response models.* It may be possible in some circumstances to use epidemiological data from case-control studies, cross-sectional surveys, prospective cohort studies (relatively rare), or other designs to estimate approximate exposure-response relations. The major statistical challenge is to separately estimate exposures and conditional probabilities of adverse responses, given exposures. Issues such as recall biases, omitted variables, uncertain model forms, etc. in Table A2 can complicate valid statistical inference of dose-response functions from epidemiological data, or even make it impossible. However, progress in advanced statistical modeling methods, such as mixture distribution modeling with an unknown number of mixture components, raises the possibility of using epidemiological data for exposure-response modeling.

Despite this range of theoretical approaches, in practice biologically motivated parametric dose-response models are the most common, and usually the best justified, models in widespread use. They are usually fit to data by a combination of MLE for point estimates and computationally intensive resampling techniques (e.g., bootstrapping algorithms) for confidence intervals and joint confidence regions for model parameters (Haas et al., 1999, Chapter 7, c.f. p. 293).

Haas et al. (1999, p. 98) state that "It has been possible to evaluate and compile a comprehensive database on microbial dose-response models." Chapter 9 of this monograph provides a compendium of dose-response data and dose-response curves, along with critical evaluations and results of validation studies, for the following:

- *Campylobacter jejuni* (based on human feeding study data)
- *Cryptosporidium parvum*
- Pathogenic *E. coli*
- *E. coli* O157:H7 (using Shigella species as a surrogate)
- *Giardia lamblia*
- nontyphoid *Salmonella* (based on human feeding study data)
- Salmonella typhosa
- *Shigella dystenteriae, S. flexneri,* etc.
- *Vibrio cholerae*
- Adenovirus 4, Cocksackie viruses, Echovirus 12, Hepatitis A virus, Poliovirus I (minor), rotavirus

Thus, for many food-borne and water-borne pathogens of interest, dose-response models and assessments of fit are already available. However, some of these existing models have parameter values estimated from data collected from limited subpopulations (e.g., healthy young male student volunteers for feeding studies). It may be necessary to modify these models for other subpopulations, e.g., by multiplying by estimated relative risk factors. On the other hand, many of the existing models have been validated using outbreak data and other epidemiological sources, and the best-fitting models (often, the Beta-Poisson model) usually compare quite favorably to available data.

METHODS OF EXPOSURE-RESPONSE MODELING TO AVOID

Only validated exposure-response models should be used for purposes of quantitative risk assessment. Unvalidated models may be used in preliminary screening , e.g., to establish whether a risk might be large enough to justify a risk management intervention. They should not be used for final decision-making.

Aggregate population-based statistical models that do not distinguish between exposure levels and exposure-response relations have sometimes been proposed. These seek to simplify the risk analysis process by eliminating the need to consider exposure-response relations, e.g., by developing an aggregate regression relation between contaminated food produced and adverse health effects in the population. In general, such approaches give inaccurate results that cannot be interpreted causally, and should be avoided. Assumption-driven statistical models (e.g., any of the empirical exposure-response relations) that have not been validated should also be avoided.

DEALING WITH UNCERTAINTY IN EXPOSURE-RESPONSE MODELS

If the correct dose-response model form is unknown and several models all provide adequate fits to the available data, then multiple plausible models may be used to carry out the rest of the assessment. In this case, the analysis can be organized and presented as a *decision tree* in which the choice of model form leads to different branches in the tree. The results of the risk analysis at the end of each branch of the tree are contingent on the assumptions and modeling choices that lead to it. Different branches may be weighted by the relative strength of the evidence supporting them. This approach can be used to present and analyze uncertainties due to choices of dose metrics, response definitions, and other modeling decisions, as well as choices of particular dose-response models.

Note: Branches in a model uncertainty decision tree can be weighted using formal statistical criteria such as the AIC, BIC, Mallows' criterion, cross-validation results, etc. These measures for model evaluation and model selection are now built into many statistical software packages, including SAS. However, in practice, it is more usual to rely on subjectively judged weights of evidence to combine results across multiple branches of the model decision tree. (See Sielken RL Jr, Bretzlaff RS, Stevenson DE. Challenges to default assumptions stimulate comprehensive realism as a new tier in quantitative cancer risk assessment. *Regul Toxicol Pharmacol.* 1995 Apr;21(2):270-80.)

Within each exposure-response model form, uncertainties about parameter values and model predictions can be quantified using computationally intensive resampling techniques to compare results based on multiple subsets of the available data. Model cross-validation and bootstrap techniques can be used to estimate the predictive power of a model and to estimate joint confidence regions for its uncertain parameter values, respectively.

VALIDATION OF EXPOSURE-RESPONSE MODELS

Dose-response models should be validated by analyzing epidemiological data, especially data from outbreaks (Haas et al., 1999, Chapters 7 and 9). The validation step allows the predictive accuracy of the models accuracy to be critically assessed.

One way to carry out dose-response model validation with outbreak data is to use the dose-response model, together with the estimated attack rates and durations of exposures and estimated quantities ingested, to predict the most likely dosages in the contaminated media that caused the outbreak, the most likely illness ratio during the outbreak, and levels of other observed quantities. The predicted levels can then be compared to the actually measured or observed values recorded during the investigation of the outbreak. See Haas et al. (1999) for details.

If the predicted quantities (e.g., contamination levels, illness ratios, etc.) do *not* match the validation data, then the exposure-response model should be corrected. This is done by refining the model to include other relevant variables and/or by using the differences between predicted and observed values to select more appropriate mathematical model forms that will explain and reduce the differences. If the predicted exposures adequately match validation data, as indicated by goodness-of-fit tests and model diagnostics (e.g., plots of residuals), then the exposure-response model may be used to make predictions for risk assessment within the validated range of conditions. In this case, remaining uncertainty in model parameters and predictions should be expressed through confidence intervals for single quantities (e.g., the mean illness rate in the population) and through joint confidence regions for multiple correlated quantities, such as the risks experienced by members of different subpopulations.

METHODS FOR RISK CHARACTERIZATION

Risk characterization is a purely arithmetic process. Given the results of the exposure assessment and exposure-response modeling steps, the risk metrics and confidence intervals can be calculated by Monte Carlo simulation uncertainty analysis applied to the analytic probability formulae such as the following:

Pr(effect E occurs in person j on day d) =

$$\Sigma_x \Pr(\text{effect E occurs in j} \mid \text{j's exposure} = x)\Pr(\text{j's exposure} = x \text{ on day d})$$

The summation is performed over different exposure levels, x. The two terms are:

- Pr(exposure of j = x on day d) = probability for j's exposure level on day d. This is obtained from the exposure model. The probabilities of different exposure levels can be conditioned on any available information about individual j. Typically, exposure is treated as a random variable, and Monte Carlo uncertainty analysis is used to repeatedly sample values of exposure from its conditional distribution. These sampled values are used to calculate corresponding probabilities of effects in the above formula.

- Pr(effect E occurs in j | j's exposure = x) is the exposure-response probability relation obtained from the exposure-response model. The response probabilities at different exposure levels can be conditioned on any available information about individual j. Typically, the dose-response model is treated as uncertain, and Monte Carlo uncertainty analysis is used to repeatedly sample values of its parameters from their joint conditional distribution, given all available data. These sampled values are used to calculate corresponding probabilities of effects in the above formula.

Repeated sampling of individuals (randomly sampled from the joint frequency distribution of individual covariates and exposures) and of exposures and dose-response relations given individual characteristics, allows all of the output risk metrics, confidence intervals, and confidence regions to be automatically calculated as accurately as desired. It is common practice to use commercial Monte Carlo uncertainty analysis software products, such as Analytica™, @RISK™ or Crystal Ball™, to automatically perform the required simulations, collect the results, and display the output risk metrics, confidence intervals, and joint confidence regions. These products use special simulation techniques (such as antithetic variate variance reduction, importance sampling, Latin Hypercube sampling, etc.) to reduce the CPU time needed to obtain accurate answers.

In special cases, risk characterization calculations can be carried out symbolically or analytically. However, the current state of practice generally relies on Monte Carlo uncertainty analysis to obtain fast, accurate numerical answers. Guidance and principles for using and documenting Monte Carlo uncertainty analysis in risk characterization are

available in: Burmaster DE, Anderson PD. Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments. Risk Anal. 1994 Aug;14(4):477-81.

METHODS OF RISK CHARACTERIZATION TO BE AVOIDED

Alternative methods of risk characterization have sometimes been proposed. For example, rather than building up risk from exposure assessment and exposure-response components summed (or integrated) over all exposure levels weighted by their respective probabilities, it has sometimes been proposed to directly estimate total risk in a population and then to use aggregate regression models or equations to attribute some part of this total to animal antimicrobial use. In general, this approach necessarily requires subjective policy judgments about attribution (since attributable risks cannot be uniquely objectively defined from data in multivariate models where different factors, such as the care taken by food handlers, processors, consumers, and farmers interact to determine total risk.) It may also produce results that are causally irrelevant, since attributable risk fractions usually do not correctly predict the changes in effects that would be caused by changes in exposures or other inputs. Therefore, these methods are not recommended.

VALIDATION OF RISK CHARACTERIZATION RESULTS

Ideally, risk characterizations should be validated if possible, although this is not strictly required if the exposure assessment and exposure-response variables have already been validated. Risk characterizations can be validated, if adequate data are available, by applying their exposure assessment models and exposure-response models to multiple distinct populations (e.g., in different geographic sub-regions and/or in different seasons or years). The predicted risks for each subpopulation are then compared to observed values of illness rates and other metrics to determine whether the observed values could plausibly have been drawn from the predicted risk distributions. Formal goodness-of-fit tests and model diagnostics are used to compare observed and predicted values.

If the predicted risks do *not* match the validation data, then model inputs, assumptions, and functions should be checked. They should be refined if necessary by using the differences between predicted and observed values to make changes in the model that will explain and reduce the differences. If a "comprehensive realism" decision tree (Sielken RL Jr, Bretzlaff RS, Stevenson DE., 1995) has been used to organize and display modeling uncertainties, then the weights of evidence for different branches may be updated to increase the relative weights on branches (i.e., assumption sets) that yield predictions that are most consistent with the validation data.

If the predicted risk metrics do adequately match validation data, as indicated by goodness-of-fit tests and model diagnostics (e.g., plots of residuals), then the underlying risk model may be used to make predictions for risk assessment within the validated range

of conditions. In this case, remaining uncertainty in model parameters and predictions should be expressed through confidence intervals for single quantities (e.g., the mean illness rate in the population) and through joint confidence regions for multiple correlated quantities, such as the risks experienced by members of different subpopulations.

METHODS AND DATA FOR UNCERTAINTY AND SENSITIVITY ANALYSIS

Methods for uncertainty analysis have already been discussed for the exposure assessment and exposure-response models. They include:

- *Monte Carlo uncertainty analysis* using commercial software products such as Analytica™, @RISK™, Crystal Ball™)

- *Bayesian uncertainty analysis* for estimation of joint confidence regions for model parameters and predictions. (e.g., using the free WINBUGS software for Markov Chain Monte Carlo Bayesian inference with missing data)

- *Bootstrapping and other resampling techniques* for estimating joint confidence regions for model parameters and predictions.

- *Model cross-validation for* estimating the accuracy and prediction error characteristics of model predictions from performance on multiple subsets of data.

Commercial software tools for uncertainty and sensitivity analysis of risk models are increasingly available and appropriate for use in antimicrobial risk assessment as well as other areas of risk analysis. See e.g.,

- http://www.palisade.com/html/risk/new_in_risk45.html
- http://www.merak.com/kr/files/316/DTreeAboutthisrelease.pdf

METHODS OF UNCERTAINTY AND SENSITIVITY ANALYSIS TO AVOID

Many past models either ignore model uncertainties or use unvalidated default assumptions and then state that the risk results are contingent on these assumptions. Given the emerging widespread availability of high-quality uncertainty and sensitivity analysis software, risk analyses should now be expected to quantify and present all key sensitivities, show estimated variability of risk metrics in the exposed population, and provide uncertainty analysis displays for their major conclusions.

METHODS AND DATA FOR RISK MANAGEMENT DECISION-MAKING

Formal methods for risk management decision-making apply the methods and frameworks of decision analysis, optimization, and group decision-making to clarify value trade-offs among competing goals and to select risk management options that correspond to the most preferred probability distributions of consequences. This can be done if relative preference weights, called utilities, can be assigned to different possible consequences of the risk management decisions being evaluated (e.g., to different values of the individual and population risk metrics). The probabilities of these different consequences for different risk management decisions are obtained from the output of the risk assessment. Each risk management decision being considered leads to a corresponding set of probabilities for different consequences and their utilities. The decision leading to the greatest mean value of the utility is recommended.

In practice, this formal decision-analytic approach is seldom directly applicable. Different participants may have different preferences for outcomes, be willing to make different trade-offs among goals (e.g., minimizing average risk vs. reducing inequities in the distribution of risks), and have different tolerances for accepting risks. In such cases, agreed-to utilities for different consequences may not exist, and risk management decision-making requires negotiation and compromise as well as analysis and deliberation. Although the formal process may not be directly applicable, its conceptual framework is still useful for organizing analysis and deliberation, separating beliefs from preferences for consequences, and identifying and resolving relevant conflicts and/or uncertainties about facts and values.

PARTICIPATORY RISK-MANAGEMENT DECISION PROCESSES

Societal risk management decisions are usually made by multiple participants and reflect the interests of multiple stakeholders with partially conflicting interests and beliefs. The participants interact through *decision processes* in which individual proposals, choices, offers, commitments, and actions or behaviors are iteratively modified until an outcome is reached. In general, *risk management decision processes* refer to procedures, typically with multiple stages or steps, by which multiple participants jointly determine how risks are to be managed. Each participant uses information about what others have done, claimed, or offered to decide what to do next. Their interacting decisions determine how risks are managed.

Properties of a risk management decision process that are often associated with its perceived legitimacy, and hence with effectiveness in changing people's attitudes and behaviors (e.g., Slovic, 1999), include the following:

- Identify and involve key players (or "stakeholders") early on whose expertise, participation, assent or consent will later be needed.

- Give each stakeholder opportunities and a positive incentive to participate (e.g., an expectation of helping to make collective choices that he prefers).
- Allow individual concerns, preferences and values to be surfaced, acknowledged, and responded to. Confront and resolve conflicts among individual beliefs and/or preferences using stated principles for how decisions should be made when individuals disagree.
- Partner with stakeholders to build trust in the process, get it used, and improve it over time.

Techniques for managing group dynamics and for organizing and running effective meetings and hearings can often create a broadly shared perception that most of these elements have been accomplished. For example, making sure that all stakeholders are given opportunity to comment; recording and systematically responding to (or at least noting) points raised; and actively encouraging participation are simple methods that go far toward making a process look and feel legitimate. Allowing participants to take turns speaking, keeping and publishing careful notes and written responses to questions and issues raised, and providing multiple opportunities to review and comment before a final decision is made are all methods for creating perceived legitimacy for public risk management processes.

All group decision processes for risk management have some intrinsic limitations. For example, those who set the agendas for group decision processes (and process the results) may be able to manipulate the probable outcomes even for decision processes (e.g., voting) that are widely perceived as fair and legitimate. If there is private information, then strategic misrepresentation of interests and beliefs may also hamper the success of decision processes in obtaining fair, efficient outcomes with high probability.

Approaching risk management decision processes as exercises in *joint problem-solving* by the participants, backed by a commitment to use mutually agreed-on principles and procedures (e.g., of fairness or voting) to resolve conflicts when necessary, provides a powerful practical approach for creating consensus and acceptance of outcomes despite these potential limitations.

VALIDATION OF RISK MANAGEMENT RESULTS

A risk assessment model predicts the probable human health effects and other consequences of different risk management actions by predicting their impacts on human exposures to microbial loads. Following implementation of a risk management decision, these predictions should be tested. This is done by conducting an *evaluation study* to assess whether the predicted changes in exposures and health effects actually occurred. If not, the risk assessment model may need to be refined (see Validation of Risk Characterization) and the recommended risk management decision may have to be revised.

APPENDIX B: Example of a Quantitative Health Risk-Benefit Model for Macrolide Use in Chickens

This Appendix develops a simple algebraic population risk-benefit model (equations (7) and (8) below) and then estimates its parameters and calculates its outputs using recent data for macrolides. The model and parameter estimates are used to support various examples in the main text.

Model Formulation

This section develops a relatively simple algebraic model of human health impacts from changes in animal antibiotic use, based on the idea that the population human health risk from animal antibiotic use, defined and expressed in units of expected numbers of adverse consequences of different severities (e.g., mild, moderate, severe, or fatal cases) (Buzby, et al., 1996) per year, can be calculated from the formula:

Population risk = (expected exposures per year) * (expected consequences per exposure). (1)

The incremental population risk caused (if positive) or prevented (if negative) by a change in animal drug use is the change in the expected adverse consequences per year, as determined by the above product. This model is appropriate for sporadic cases of foodborne illnesses, such as campylobacteriosis, that are well approximated by Poisson processes, for which the expected value determines the entire probability distribution of values (FDA, 2001). Using conditional probabilities, it can also be expanded as:

Population risk = (expected exposures per year) * (expected illnesses per exposure) * (expected adverse consequences per exposure) (1a)

in order to take advantage of separate information about illness rates and their clinical consequences, such as number of illness days by severity category.

For a risk management action that affects multiple types or pathways of exposure (e.g., both susceptible and resistant strains of one or more pathogens, perhaps transmitted via more than one food animal commodities, or by chicken meats from both AS+ (airsacculitis-positive) and AS- flocks) and/or that affects multiple human subpopulations having significantly different exposure and/or dose-response characteristics, the above "risk = exposures * consequences" formulas can be applied to each exposure path and subpopulation, and the results summed to obtain total population risk. (This is because of the additivity of Poisson processes. Statistical algorithms to determine significantly different subpopulations from multivariate data include nonparametric clustering and partitioning methods such as classification tree analysis (Cox, 2002).) For example, the formula (1) for individual risk (expressed as expected illnesses per capita-year from chicken consumption) if a fraction F of chicken servings come from carcasses from AS+ flocks while the rest are from AS- flocks, would become:

E(illnesses per capita-year) = [E(illnesses | eat serving from AS+ chicken)*F + E(illnesses | eat serving from AS- chicken)(1 – F)]*(Number of chicken servings eaten per year) (2)

Here, E(illnesses | eat serving from AS+ chicken), for example, is the probability of a person becoming ill from eating a serving from a chicken from an AS+ flock, if illnesses are not

transmitted among people. (If one ill person might infect others, then this probability is replaced by the expected number of illnesses caused per serving.) Equation (2) may be abbreviated as:

$$\Pr(\text{illness}) = [\Pr(\text{illness} \mid \text{AS+})*F + \Pr(\text{illness} \mid \text{AS-})*(1 - F)]*M \quad (2a),$$

where F = Pr(AS+) is the probability that a randomly selected serving is from an AS+ bird; and M = number of chicken servings consumed per year.

The same formula can be extended to sum over different types or severities of illnesses. For example, it can be applied as follows to calculate expected quality-adjusted life-years (QALYs) lost per capita-year from chicken servings that may be contaminated by both susceptible and resistant *C. jejuni*:

E(QALYs lost per year from chicken consumption) = {E(QALYs lost | susceptible illness)*[Pr(susceptible illness | AS+)*F + Pr(susceptible illness | AS-)*(1 – F)] + E(QALYs lost | resistant illness)*[Pr(resistant illness | AS+)*F + Pr(resistant illness | AS-)*(1 – F)]}*M (3).

If the conditional probability of an illness being antibiotic-susceptible, given that one occurs, is approximately the same for cases from AS+ and AS- birds, then we will denote it by:

s = Pr(susceptible illness | illness) = fraction of chicken-caused campylobacteriosis cases that are antibiotic-susceptible.

(If s is different for AS+ and AS- flocks, then separate s^+ and s^- values can be propagated throughout the analysis, but there is currently no empirical need for such generality, so we will simply use a single s.) Introducing the shorter notations: P^+ for Pr(illness | AS+) (or, more generally, P^+ = E(illnesses | AS+), if ingesting a single contaminated serving from an AS+ bird might lead to multiple illnesses in one or more victims, e.g., due to contagious infection or increased vulnerability to further illnesses); P^- for Pr(illness | AS-) (or, more generally, P^- = E(illnesseses | AS-)); Q_s = E(QALYs lost per susceptible illness); and Q_r = E(QALYs lost per resistant illness), the above formula can then be written as:

$$\text{E(QALYs lost per capita-year from chicken consumption)} = [sQ_s + Q_r(1 - s)][(P^+)F + (P^-)(1 - F)]M$$

$$= [Q_r + s(Q_s - Q_r)][P^- + F(P^+ - P^-)]M \quad (4)$$

This still has a simple logical form: "Risk = (illnesses-generating exposures per capita-year)*(health consequences per illness)", where these two terms have been expanded more explicitly as:

- Expected illness-generating exposures per capita-year = $[FP^+ + (1 - F)P^-]M$
- Expected health consequences per illness = $[sQ_s + Q_r(1 - s)]$

If QALYs are not desired as an aggregate summary measure of harm, then these formulas can be adapted to estimate expected illness-days or number of illnesses by severity category. For example, suppose that a regulator wanted only to know the expected *number* of antibiotic-resistant cases of campylobacteriosis each year that are treated with that antibiotic, without regard for the clinical consequences of such a treatment (as in FDA, 2001). Then the correct calculation formula would be equation (4) with the QALY weights changed to $Q_r = 1$ and $Q_s = 0$. If the regulator cared only about the total number of illnesses, then equation (4) with $Q_r = 1$ and $Q_s = 1$ would give this number.

Now, suppose that the regulator wants to know the human health consequences of a proposed ban on animal antibiotic use that is expected to prevent a fraction p of all current chicken-borne antibiotic-resistant illnesses, replacing them with antibiotic-susceptible illnesses instead. (p is the *preventable resistance fraction* for the ban. It is intended to have an explicit causal interpretation, unlike attributable fractions (Cox, 2001, Chapter 4).) If the proposed ban would also cause an incremental fraction ΔF of currently AS- flocks to become AS+, then the risk after the ban would be:

$$E(\text{QALYs lost per capita-year}) = \{[p(1 - s) + s]Q_s + Q_r(1 - s)(1 - p)\}[(P^+)(F + \Delta F) + (P^-)(1 - F - \Delta F)]M = [Q_r + (p - ps + s)(Q_s - Q_r)][P^- + (F + \Delta F)(P^+ - P^-)]M$$

Subtracting the pre-ban risk (given by this same formula with $p = \Delta F = 0$) and simplifying yields:

$$E(\text{change in QALYs lost per capita-year if ban is implemented}) = \{[Q_r + s(Q_s - Q_r)][\Delta F(P^+ - P^-)] + p(1 - s)(Q_s - Q_r)[P^- + F(P^+ - P^-)]\}M$$

For a population of N identical individuals, the total population health impact from the ban would be:

$$E(\text{change in QALYs lost per year caused by ban}) = \{[Q_r + s(Q_s - Q_r)][\Delta F(P^+ - P^-)] + p(1 - s)(Q_s - Q_r)[P^- + F(P^+ - P^-)]\}MN \quad (5)$$

Again, the weights Q_s and Q_r may be adjusted to calculate various quantities of interest, such as expected changes in total number of illnesses per year (set $Q_r = Q_s = 1$) if the ban is implemented.

While equation (5) may perhaps not seem close to the goal of a clear, simple risk formula, a substantial simplification holds for *C. jejuni* infections in the US. At present, airsacculitis rates in chicken flocks are so low that their contribution to observed campylobacteriosis rates is negligible and the fraction of AS+ flocks can be approximated as $F \approx 0$. Using this approximation, the model for the human health effect of a ban on antibiotic use affecting airsacculitis simplifies to:

$$E(\text{change in QALYs lost per year caused by ban} \mid F = 0) = \{[Q_r + s(Q_s - Q_r)][\Delta F(P^+ - P^-)] + p(1 - s)(Q_s - Q_r)(P^-)\}MN \quad (6)$$

This is easily interpretable. The term $p(1 - s)(Q_s - Q_r)(P^-)MN$ corresponds to: the expected number of total current number of *C. jejuni* illnesses per year in the population (MNP^-) times the fraction that are resistant, $(1 - s)$, times the fraction of these resistant illnesses that would be eliminated by a ban and replaced by susceptible illnesses (p), times the change in health impact for each such case, $(Q_s - Q_r)$. In other words, it is just the *expected direct human health benefit* from the ban due to fewer resistant (and more susceptible) illnesses being created per chicken serving ingested. This may also be defined as the preventable *human health risk* from continued use of animal antibiotics under current conditions. We refer to this component of equation (6) as a top-down or *farm-to-clinic model* (in contrast to a bottom-up farm-to-fork model) as it begins with total current number of *C. jejuni* illnesses per year in the population and then apportions a fraction of this total to the contribution from animal antibiotic use on the farm. $(Q_s - Q_r)$ is typically negative (i.e., the change in health impact from reduced resistance is a reduction in days lost. To present the same quantity with a positive sign, we may interpret $p(1 - s)(Q_r - Q_s)(P^-)MN$ as the *expected illness-days prevented per year* by a ban. Finally, if illness days are prevented only for a fraction f*r of cases that (a) Are treated with a resisted drug (fraction = r); and (b) experience treatment failure due to resistance to that drug (fraction = f, i.e., treatment would have been uncompromised for $(1 - f)$ of

treated resistant cases despite the resistance), then the expected number of illness-days per year prevented by a ban becomes:

$$\text{Expected illness-days prevented per year} = p(1 - s)fr(Q_r - Q_s)(P^-)MN \quad (7)$$

On the other hand, the term:

$$\text{Expected illness-days caused} = [Q_r + s(Q_s - Q_r)][\Delta F(P^+ - P^-)]MN \quad (8)$$

also has a simple intuitive interpretation. It represents the fractional increase in servings per year from AS+ flocks instead of AS- flocks, (ΔFMN), times the change in risk of illness from each such serving, ($P^+ - P^-$), times the average health impact of each illness, $[Q_r + s(Q_s - Q_r)]$. In other words, it is the preventable *human health loss (i.e., risk) from the ban*. This is also the expected human health benefit from continued animal antibiotic use under current conditions.

Equation (7) constitutes a relatively simple, interpretable model for the potential *net human health impact of changing the status quo to cease animal antibiotic use*. Additional refinements could be made to reflect the *timing* of the gradual adjustment from pre-ban to post-ban conditions, changes in parameter values over time, and so forth. However, a simple static comparison of potential risks to potential bans from a ban may be very useful in addressing the key policy-relevant question suggested by the farm-to-fork model: Which is greater, the increase in risk-per-serving from additional AS+ flocks, $[Q_r + s(Q_s - Q_r)][\Delta F(P^+ - P^-)]$, if a ban is implemented; or the reduction in risk-per-serving if a ban is implemented from reduced load of resistant bacteria, $p(1 - s)fr(Q_r - Q_s)(P^-)$? To answer this question empirically with the help of equation (6), it is necessary to estimate the model parameters from data.

Estimating Model Parameters from Data

We begin by estimating the parameters in the expression for human health benefits of a ban [i.e., preventable human health risks of continued antibiotic use = $p(1 - s)fr(Q_r - Q_s)(P^-)(MN)$] from recent data.

The model parameters (Q_r, Q_s, s, ΔF, P+, P-, p, f, r, M, N) may be significantly different for people in different geographic regions or with different ethnic and demographic attributes. Thus, in principle, individual risk models can be created for each group with a distinct combination of parameter values (corresponding to a cluster within the population or a leaf node in a statistical classification tree) and the risk for each group, weighted by its size, can be summed to estimate population risk. However, given the aggregate data available from which to estimate model parameters, we focus on estimating the average individual risk and total population risk, i.e., the average individual risk times the population size. This neglects the individual-level uncertainty and variability modeling made possible by stochastic simulation models, but allows population risk estimates to be derived in a relatively simple, easily verified way from available data.

Estimating ($Q_r - Q_s$): Human health impact of resistance

To estimate a human health impact from antibiotic use in animals, there must be a measurable or assumed difference in the human health consequences of antibiotic-resistant versus antibiotic-susceptible strains of Campylobacter, corresponding to the difference ($Q_r - Q_s$) in the model. This difference typically depends on the medical treatment that a patient receives. We distinguish among the following possible cases:

- *Patients who receive no treatment:* In this case, we assume that there is *no difference* in average medical outcomes, i.e., $Q_s = Q_r$. Barza and Travers (2002) suggest that resistant *C. jejuni* may opportunistically infect a patient treated with antibiotics for some other reason. This seems plausible *a priori*, unless antibiotic treatment *in vivo* is effective against (*in vitro*) resistant bacteria as well as susceptible bacteria. However empirically, in recent case-control data (e.g., Friedman et al., 2000, Effler et al., 2001, raw data analyzed in Cox, 2002), patients who eat chicken and take antibiotics for non-campylobacteriosis reasons are not at greater risk of resistant campylobacteriosis than patients who do not. Indeed, in these data sets, chicken consumption is associated with significantly *smaller* risks of campylobacteriosis [perhaps due to acquired immunity, previously demonstrated for raw milk consumption (Blaser et al., 1987; see also Walz et al, 2001)].
- *Patients who receive treatment but don't need it.* Again, we assume that, in these patients, that there is *no difference* in average medical outcomes, i.e., $Q_s = Q_r$. In agreement with Ang and Nacham, 2003 (*op cit*), the CDC states that:

> "The disease is usually self-limiting, so antibiotic treatment is only indicated in severe cases. Erythromycin is the drug of choice, with ciprofloxacin as a suitable alternative in adults. Cases of septicaemia are best treated with gentamicin, but erythromycin, chloramphenicol and tetracycline may also be used. The faeces often remain positive for 2-7 weeks, but long-term carriage is rare. Treatment with erythromycin may significantly shorten the duration of excretion." (http://www.cdc.gov/ncidod/dbmd/diseaseinfo/campylobacter_g.htm)

  Thus, *only severe cases of campylobacteriosis warrant treatment* with antibiotics. In clinical practice, when a diagnosis of campylobacteriosis has not yet been made fluoroquinolones or other broad spectrum antibiotics may be prescribed as empiric treatments, so that many non-severe cases end up being prescribed these antibiotics that would not be if better information were available. However, this practice brings no clear clinical benefit to adult patients and is changing as more rapid, accurate diagnostic tests for campylobacteriosis become available (e.g., Endtz et al., 2000) and as physicians become more sensitive to the importance of halting over-prescriptions of human antibiotics in non-essential cases where no clinical benefit is expected (WHO, 2003). We assume that there are *no human health impacts,* either beneficial or adverse, from prescribing antibiotics to non-severe cases for whom treatment would not be indicated if the physician had perfect diagnostic information.
- *Patients who receive treatment and for whom treatment is appropriate.* Current clinical practice, as reviewed above, specifies that treatment with erythromycin or ciprofloxacin may be indicated for severe diagnosed cases. We focus on these cases as the ones for which appropriate treatment with antibiotics is recommended to achieve potential clinical benefits. However, the average true clinical benefit from treatment in terms of resolution of symptoms remains controversial (Ang and Nacham, 2003, *op cit*). The extent to which *severe* cases of *C. jejuni* (which are often associated with AIDS or other severe underlying illnesses and have symptoms that have lasted for over a week) receive erythromycin and ciprofloxacin treatment, as well as the relative clinical impacts of those treatments in resistant and susceptible cases, are not known. Severe cases comprise only less than 1% (approximately 0.595%) of all cases of campylobacteriosis (Buzby, et al., 1996), making the study of clinical outcomes specifically for severe cases difficult for existing case-control data. However, as a baseline value, and in the absence of more relevant (specifically chicken-associated, severe illness) data, we will assume that a plausible worst case is that treatment failures might create an average of two excess days of treatment as physicians monitor the results of the treatment and/or wait for the results of a resistance test before switching to an alternative therapy (Ang and Nacham, 2003, *op cit*). Thus, we make the baseline assumption that $(Q_r - Q_s) = 2$ days of excess illness and treatment for severe cases that fail to respond normally to erythromycin therapy due to resistance. Sensitivity analysis is then used to study the effects of varying this starting assumption.

The fraction f of patients with severe campylobacteriosis who fail to respond normally to initial

erythromycin therapy due to resistance is not known. To be conservative, we assume that the true fraction could be as high as 1.

To bound uncertain quantities by their most extreme possible values, we assume that *all* severe cases of campylobacteriosis seek treatment and that *all* are prescribed antibiotics. According to CDC-FoodNet surveillance data, about 55% of campylobacteriosis patients who are prescribed antibiotics receive fluoroquinolones (usually, ciprofloxacin) (FDA-CVM, 2001). Assuming that fewer than 20% of those fluoroquinolone-treated cases eventually switch to erythromycin or another macrolide, the total fraction of severe campylobacteriosis cases treated with macrolides is at most about 0.50.

Estimating M, N, and P^-: Total severe C. jejuni cases per year from chickens in the US

The first six lines of Table 4 summarize the parameters and data sources used to estimate the number of severe campylobacteriosis cases per year caused by consumption of food contaminated by chicken-borne *C. jejuni*, corresponding to the product $(P^-)(MN)$ in the model when the model is applied to severe cases only. (Henceforth, all calculations are made only for the treatment group consisting of patients with severe campylobacteriosis who receive antibiotic treatments. Although the model formally requires performing calculations for the other treatment groups, i.e., persons without severe cases and/or without antibiotic treatment, and then summing the results over all groups to calculate population risk, in practice, as explained above, the health impact parameter $(Q_s - Q_r)$ is assumed to be 0 for these other groups, making it unnecessary to carry out the calculations for them.)

Since essentially all chicken-borne *C. jejuni* cases are assumed currently to come from AS-flocks, the current average risk of campylobacteriosis per serving of chicken from AS- flocks (including possible effects of cross-contamination to other foods in the kitchen), denoted by P^- in the model, can be estimated by dividing the estimated total number of chicken-caused (severe) campylobacteriosis cases per year by the total estimated number of chicken servings ingested per year:

P^- = (total chicken-caused severe cases)/(total servings) = (total severe cases * fraction from chicken)/(MN).

The two quantities M and N are readily available: N = number of people in the US ≈ **292E6** (US Census), while M = average chicken servings per capita-year in the US has been estimated as **38.0 servings/year** for "fresh" chicken that might be carrying *C. jejuni* (Cox and Popken, 2002, based in part on data from FDA-CVM, 2001). Precision in these estimates is unnecessary as uncertainty about these two parameter values cancels out in calculations of relative risks and relative risk:benefit ratios for a ban: only their product (MN) is used in the model (equation (6)) to scale from relative to absolute numbers of cases with and without a ban.

Multiplying the first five factors from Table 4 gives the estimated total number of severe *C. jejuni* cases per year in the US:

(13.37 reported campylobacteriosis cases/100,000 capita-year)*(0.00595 fraction that are severe enough for antibiotic treatment to be indicated)*(8 assumed severe cases per reported case)*(292,000,000 people in US)*(0.99 *C. jejuni*) = (13.37)*(0.00595)*(8)*(2920)*(0.99) = **1,840** severe *C. jejuni* cases per year.

Uncertainties in this calculation include uncertainty about the true proportions of *C. coli* and other non-*jejuni* spp. (assumed here to be only 1% based on general CDC data, but possibly on the order of a few percent) and considerable uncertainty about the true number of cases for each reported case. Mead et al. (1999) (*op cit*) suggested 38 as a reasonable guess for most (non-severe) cases, in the absence of Campylobacter-specific information, and "arbitrarily used a far lower multiplier of 2" for severe cases. Since we focus on severe cases, 2 may be more appropriate as an under-reporting factor than 38. Rather than select a single number for this uncertain quantity, we will treat it as a subjectively random variable with a plausible range (interpreted as a subjective 95% probability interval) from 2 to 38. This subjective uncertainty assessment can be expressed as a random variable with a median value of **8** with an uncertainty factor of about 5 (i.e., the true value could be from 8/5 to 8*5).

Such geometric medians and uncertainty factors for highly uncertain parameters values are useful for quickly approximating and expressing uncertainties in cases where the published literature specifies a wide range of possible values for the same uncertain quantity. They also support approximately log-normal uncertainty distributions for the outputs of sequences or networks of calculations with multiple uncertain quantities (Druzdzel, 1994). If each term in the benefit formula $p(1 - s)(Q_s - Q_r)(P^-)(MN)$ is characterized by a point estimate and a multiplicative uncertainty factor, interpreted as a geometric mean (or median) and geometric 95% confidence interval, then these estimates can easily be combined, as discussed below, to estimate the approximate log-normal distribution for the entire product.

The next factor in Table 4 is the estimated fraction of *C. jejuni* cases caused by contaminated chicken (possibly via cross-contamination of other foods). This is a crucial parameter, and, unlike the product (MN), it does not cancel out in further analyses. Uncertainties about it propagate directly to uncertainties about P^- and about the potential human health benefits from a ban on antibiotics used in chickens. Therefore, it is worth carefully examining possible empirical data for estimating this fraction.

While Mead et al. (*op cit*) assumed that 80% of campylobacteriosis illnesses are food-borne, based on a 1992 study (and translating "most" as 80%) and FDA-CVM (2001) assumed that about 57% of all *C. jejuni* cases are caused by consumption of chickens based on pre-1985 data, more recent evidence suggests that much lower numbers may be appropriate now. Based on post-2000 data for outbreaks and sporadic cases, population risks might be allocated among competing sources roughly as follows: Foreign travel: > 10% (Friedman et al., 2000); Drinking undisinfected water: > 50% (Kapperud et al., 2003 for Norway. Rates of drinking unprocessed ground water are probably lower in the US, but water-borne cases may still account for many sporadic cases.); Contact with infected pets and/or farm animals or farm visits: > 5% (Gillespie et al., 2003; Friedman et al., 2000; 18% for poultry husbandry in rural populations estimated by Potter et al., 2003); AIDS, sexual transmission: 4% or more, especially for resistant Campylobacter (Gaudreau and Michaud, 2003; Sorvillo et al., 1991); Unpasteurized milk: At least 1% (Gillespie et al., 2003). Then only the remaining 30% of cases might be due to food-borne *C. jejuni*. This would then be an upper bound on the fraction due to chicken-borne *C. jejuni*.

Genetic typing data suggest that the true fraction of campylobacteriosis cases caused by chicken consumption may be far less now than implied by previous assumptions. For example:

- Nadeau et al. (2002) found that only "**approximately 20%** of human Campylobacter isolates were genetically related to genotypes found in poultry "
- Hein et al. (2003) noted that "A small number of human isolates [**11 out of 101**] shared PFGE/AFLP types with poultry isolates [sampled at slaughter in Austria], however, further studies should also focus on the identification of other sources of *C. jejuni* infection in humans."

- Moore et al. (2003) stated that "Human campylobacteriosis is currently the most common cause of acute bacterial gastroenteritis on the island of Ireland... It was the aim of this study to examine the phenotypic and genotypic relatedness of campylobacters isolated from chickens and humans locally. Sixty isolates were subtyped using phenotyping techniques (biotyping, phage-typing), as well as genotyping techniques (multilocus enzyme electrophoresis (MEE), ribotyping) and the data compared. The frequency of shared phenotypes and genotypes between poultry and humans varied depending on the typing technique employed ranging from 98.2% of human isolates sharing a similar resistotyping (MAST) disc type with poultry strains to **20% similarity with MEE typing.**"

Interpreting such genotype data in terms of sources can be difficult and controversial (Schouls et al. (2003)). The 20% number from Nadeau et al. is about double the estimate of 10% from the genetic data of Wu et al. (2002) in a study of quinolone-resistant Campylobacter in Taiwan. Even if 20% is viewed as a plausible constraint on the maximum fraction of human isolates that could come from eating chicken, an unknown part of it may be due to common environmental reservoirs (such as contaminated water) shared by chickens, humans, dogs, lambs, and other species having overlapping *C. jejuni* genotypes.

Epidemiological data can complement genotyping data in estimating the true fraction of campylobacteriosis cases that are likely to be caused by eating chicken. For example:

- A recent prospective case-control study from Quebec (Michaud et al., 2002) identifies poultry as the "principal suspected source of infection" in only about **10%** of cases, comparable to drinking tap water at home (9%).
- Our analysis of data of Kapperud et al. (2003) from Norway (personal correspondence, not shown) suggests that 7/211 = **3.3%** of cases can be associated with eating undercooked poultry, after adjusting for other variables in a non-parametric classification tree model. (Eating undercooked poultry is associated with eating other undercooked meats, so not all the excess risk associated with eating undercooked poultry is necessarily caused by it.)
- In the Friedman et al. (2000b) CDC case-control data set, the population-attributable risk (PAR) for chicken consumption as a whole among non-travel, non-treatment-related cases is negative (an apparent "protective effect" if interpreted causally, perhaps due to acquired immunity). The PAR for eating chicken specifically in restaurants, where many meats are known to be risk factors (*ibid*), is only **3.1%** using a standard univariate PAR formula for 2 x 2 tables (http://watson.hgen.pitt.edu/~dweeks/odds_ratio.html) with input values of a = 665, b=341, c=1439, d=976.) In multivariate analyses, it is not significantly different from zero. Similarly, Effler et al. (2001), Table 1, shows that overall chicken consumption is associated with a statistically significant *lower* risk of campylobacteriosis (relative risk: RR = 0.6). Restaurant chicken is again identified as a risk factor (adjusted odds ratio = 1.8, p = 0.03) as found for other foods eaten in restaurants and commercial settings (Friedman et al., 2000b).
- An admittedly simplistic alternative approach would be to note that Campylobacter levels in processed broiler carcasses may already have been reduced by about 90% since the mid-nineties (Stern and Robach, 2003). Then, assuming a proportional reduction in human risk of chicken-borne campylobacteriosis, the true fraction of campylobacteriosis cases caused by eating chicken may have fallen from a pre-1995 value of at most 100% to a current value of at most **10%.**

Based on these data, a plausible range of values for the true but unknown fraction of campylobacteriosis cases currently caused by eating chicken might correspond roughly to a point estimate of about **0.10** with an uncertainty factor of about 3 (i.e., from 0.03 to 0.30). Because this parameter is so critical, we will also consider a sensitivity analysis in which the uncertainty factor is increased to 10 (corresponding to a range from 0.01 to 1.00) to determine how this affects confidence in the conclusions of the analysis.

The average probability that a serving of chicken from an AS- flock will cause a (severe) case of *C. jejuni* can be estimated from the above numbers. It is:

P^- = (total severe cases * fraction from chicken)/(MN) = (1840*0.10)/(38*292,000,000) = **1.6583E-8** average severe *C. jejuni* campylobacteriosis cases caused per chicken serving (from an AS- flock).

This estimate of P^- is based on population averages, without accounting for interindividual variability (including variations associated with ethnic and demographic attributes) in model parameters such as numbers of meals eaten, thoroughness of cooking, differences in immune status and vulnerability, etc. Thus, it should only be used for calculating aggregate population risks rather than risks for any specific individual. The expression $(P^-)(MN)$ = expected number of chicken-caused severe *C. jejuni* cases per year in the US population is just the numerator, (1840*0.10) = **184 cases per year**.

Estimating ($P^+ - P^-$): Excess risk of C. jejuni cases per serving from AS+ chickens

Mortality due to *Mycoplasma gallisepticum* in broilers ranges from "low" in uncomplicated cases to 30% in complicated outbreaks, especially during cold months (Yoder, 1991). Most birds in infected flocks survive, but their carcasses are often underweight and contribute disproportionately to increased fecal contamination and microbial loads during processing. In a recent study, airsacculitis-positive (AS+) flocks were associated with significantly increased average levels and incidence of *Campylobacter, E. coli,* and *Salmonella* at processing (Russell, 2003). AS+ flocks have greater variability in carcass sizes (see also Engster, et al., 2002) and weakened digestive tracts, which in turn increase processing errors and increase fecal contamination levels and microbial loads. Russell measured microbial *Campylobacter* spp. loads (cfu/ml) on carcasses of AS+ and AS negative (AS-) flocks before the inside/outside bird wash (IOBW) step of chicken processing. The mean $\log_{10}$ microbial load of campylobacter colony-forming units (cfus) for AS+ flocks was 1.09 while the mean for AS- flocks was 2.09; thus, the microbial load was 10-fold higher for the AS+ flocks. Although there was considerable flock-to-flock variability, this ten-fold increase in Campylobacter loads could have significant human health consequences when averaged over multiple AS+ flocks.

To estimate the corresponding risk from chicken servings from AS+ flocks, a refined exposure model is needed to account for effects of airsacculitis (AS+) on microbial loads of *C. jejuni*. To this end, we use a model previously suggested by FDA's CVM (www.fda.gov/cvm/antimicrobial/RRAIntro.pdf) for risk assessment of campylobacter. The model assumes that:

- The $\log_{10}$ of the microbial load distribution of campylobacter reaching consumers via chicken servings is approximated by an exponential distribution. CVM suggested this distribution for modeling variability in chicken-borne campylobacter exposures based on mathematical convenience rather than on empirical grounds. It may be plausible insofar as it places greater probability densities on smaller microbial loads over several orders of magnitude of cfu/ml.
- Microbial load distributions at the point of consumption are proportional to microbial loads following processing. An uncertain reduction factor a expresses proportionality between cfu/ml measured in processing rinse fluids and cfu/chicken serving at the point of ingestion.

These assumptions imply the mathematical model:

Pr(log ingested dose > x | AS-) = $e^{-a?x}$,

where

1/? = mean number of log-cfu/ml of chicken rinse fluid at post-processing (= ln(10^1.09) = 2.51 based on data of Russell, 2003, implying an estimated value for ? of ? = 1/2.51 = 0.3984.)

Substituting log(MID) for x then gives a formula for Pr(ingested dose > MID | AS-) = Pr[log(ingested dose) > log(MID) | AS-] = $e^{-a?\log(MID)}$, where MID = minimum infectious dose (possibly 1 cfu). Independently, we can also estimate the probability of an infectious dose in a chicken serving under current conditions (i.e., Pr(ingested dose > MID | AS-)) from data, as follows:

- *Main formula:* Pr(ingested dose in a chicken serving > MID | AS-) = (total cases per year from chicken consumption)/(total chicken servings per year * fraction of infectious servings that cause cases of illness)
- *Total cases per year from chicken consumption* = (1.337E-4 reported cases per capita-year, from CDC, 2003)*[38 estimated total cases per reported case (all cases, not just severe ones treated with macrolides), from Mead et al., 1999]*(0.10 estimated fraction of cases from chicken consumption, from above based on genotyping, epidemiological, and historical (Stern and Robach, 2003) data)*(292E6 people in US, from U S Census data) = (1.337E-4)*38*0.10*292000000 = 148,354 cases/year. If 99% are *C. jejuni*, then this can be rounded to **1.47E5** *C. jejuni* cases per year from chicken consumption. In this calculation, *all* cases are estimated (rather than only severe ones warranting antibiotic treatment and resistant to such treatment), as the purpose is to estimate the *total* excess cases per year (not just resistant ones or those leading to treatment failure) that could be caused by increased airsacculitis following a ban of animal antibiotics.
- *Total chicken servings per year* = (M = 38 fresh chicken servings/capita-year)*(N = 292E6, from above)
- *Fraction of infectious servings that cause illnesses* was estimated by Rosenquist et al., 2003 as about 0.2 based on human feeding data and assuming, as in WHO, 2002, that the conditional probability of illness *given* an infectious dose is approximately constant, independent of the size of the dose. Rosenquist et al., 2003 uses an experimental value of 0.22 (11/50) with a beta uncertainty distribution. WHO, 2002 states: "In the case of the feeding trial data for *C. jejuni* A3249 the probability of illness decreases with increasing dose and as such a decreasing hazard function has been estimated (Teunis et al. (1999)). However, when the data for both strains are pooled the conditional probability of illness following infection does not exhibit a dose relationship but rather is randomly distributed (Figure 4.8). It may be appropriate in this case to use a dose independent ratio to estimate the conditional probability of illness. The conditional probability can be estimated from the feeding trial data. For A3249, out of 50 people that got infected at various doses, 11 got sick (22%), while for 81-176, out of 39 people that got infected at different doses, 18 got sick (46%). Overall, pooling all the data, a total of 29 people got sick out of 89 individuals that were infected (33%)." For comparison, Finch and Blake (1985) report a median attack rate of **0.41** in outbreaks following high exposures in various food vehicles. We will use the observed median attack rate value of **0.41** (Finch and Blake, 1985) as it is based on outbreak data from a mix of populations, strains, and food vehicles under real exposure conditions. This value is slightly higher than the 33% estimated by WHO from the feeding data.
- Substituting the above values into the main formula gives: **Pr(ingested dose > MID | AS-)** = (total cases per year from chicken consumption)/(total chicken servings per year * fraction of infectious servings that cause cases of illness) = (1.47E5 cases/year)/(38*292000000*0.41) = **3.23E-5.**

Equating this empirical estimate to the above formulas gives: